

The role of transcription factors in somatic cell nuclear reprogramming by eggs and oocytes



Ming-Hsuan Wen

Magdalene College

University of Cambridge

Dec 2018

This dissertation is submitted for the degree of Doctor of Philosophy

Preface

The work in this thesis was carried out in the Wellcome Trust/Cancer Research UK Gurdon Institute, Cambridge.

This thesis is the result of my own work. The results presented in Chapter 5 and 6 are the outcome of work done in collaboration with Angela Simeone.

I declare that none of the work presented here has previously been submitted to any other University.

It does not exceed the prescribed word limit for the relevant Degree Committee.

Summary

Somatic cell nuclear reprogramming (SCNR) by eggs is a way to forcibly transform the nuclei of terminally differentiated somatic cells to an embryonic state and gain totipotency (Gurdon et al., 1958). Additionally, induced pluripotency is applied to transform identities of somatic cells to induced pluripotent stem cells by overexpression of combinatorial Yamanaka factors (iPS, Takahashi et al., 2006). Although both approaches aim to derive cells with highest plasticity, the mechanisms and differences between these procedures are not yet clear.

In my thesis, I used quantitative polymerase chain reaction (QPCR) and RNA-seq plus 5-bromouridine 5'-triphosphate (BrUTP) pulldown to evaluate the transcriptional reprogramming by maternal factors and overexpressed transcription factors during SCNR by *Xenopus* oocytes, which are inactive in DNA replication and cell division.

QPCR measures changes in the steady-state levels of transcripts within 2 days of nuclear transfer to *Xenopus* oocytes (Oocyte-NT). Three pairs of Yamanaka factor homologs were tested by QPCR and Yamanaka factor homologues regulated similar sets of pluripotency genes in mouse embryonic fibroblasts (MEFs).

Pioneer factor mFoxA1 could not up-regulate most pluripotency genes and their binding targets of neurogenic genes in MEFs while pioneer factors are proposed to bind to their targets even if they may reside in inaccessible

chromatin. This shows that the existence of other factors is needed at specified developmental stages. Hence, gene activation by transcription factors in the Oocyte-NT system requires not only corresponding binding on regulatory elements of linked genes but transcription cooperators to exert effective gene activation.

Additionally, RNA-seq plus BrUTP pulldown measures the extent to which oocytes change the transcriptional activity of nuclei transplanted to oocytes. Through RNA-seq plus BrUTP pulldown, I compared the reprogrammed transcriptomes of embryonic and somatic cells, including mouse embryonic stem cells, mouse embryonic fibroblasts and mouse myoblasts, to demonstrate the effects of maternal factors and overexpression of transcription factors on gene activities during SCNR by oocytes.

Importantly, I find that maternal factors of oocytes and the overexpression of transcription factors exert different strategies to reprogram somatic cells. Oocyte factors reprogram the donor cell nuclei to an oocyte-steady state except for the SCNR resistance genes and *xklf2*-HA overexpression enhances expression of reprogrammable genes and activates SCNR resistance genes.

Acknowledgement

At the beginning, I thought PhD study was just to train myself as an independent and intellectual scientist. Fortunately, from this life-changing journey, I have learnt much more than what I expected and I would like to thank people as follows.

I am greatly thankful to Professor John Gurdon for leading me with critical questions, teaching me how to answer questions through countless discussions and demonstrating the way to be a professional researcher with persistent hard work.

I am appreciative to the present and former colleagues in the Gurdon laboratory for providing an environment that stimulate me to brainstorm about my projects and for the help with experiment designs and techniques. Special thanks to Dr. Jerome Jullien, Dr. Magda Koziol, Dr. Angela Simeone, Dr. Clare Pacini, Dr. Kei Miyamoto, Dilly Bradford and Nigel Garrett.

I am indebted to the Ministry of Education in Taiwan for the generous funding and to Gurdon Institute for maintaining the institute a pleasant place to work. Special thanks to all the members of Admin and medium team.

I would like to express my deep gratitude to the Department of Zoology for all the educational support and to Magdalene College for providing housing and every life-related assistant.

I would like to thank my family in Taiwan for the endless care and unconditional love and to my friends in Cambridge for your warm companionship when I met difficulties and for the countless outside-the-box conversations that have broadened my horizons. Special thanks to Dr. Oi Ying Hung, Dr. Shih-Ting Wang, Dr. Li-Ching Wang, Dr. Yen-Chun Lin, Dr. Wei-Yun Chung, Dr. Su-Xue Zhang, Amy Gunton, Jasmine Luk, Chun-Fang Yang and friends in Cambridge Taiwanese Society.

Abbreviations and explanations

9.2ng/0ng	=(relative expression of <i>xklf2</i> -HA mRNA samples)/(relative expression of no mRNA injection samples)
BrUTP	5-Bromouridine 5'-triphosphate
cKO	conditional knockout
CPM	counts per million mapped reads
DE	differentially expression/differentially expressed
Egg-NT	nuclear transfer to eggs
EpiSC	epiblast-derived stem cell
ESC	embryonic stem cell
FC	fold change
FPKM	fragments per kilobase of exon per million mapped reads
GV	germinal vesicle
HA	human influenza hemagglutinin (used as a tag for Western blotting)
hKLF4	human kruppel-like factor 4
hOCT4	human octamer-binding transcription factor (aka human POU5F1)
iPSC	induced pluripotent stem cell
$-\log_{10}(p\text{-value})$	-(the logarithm to base 10 of probability value)
MBT	midblastula transition
MDS	multidimensional scaling (a means of visualizing the level of similarity of a individual cases of a dataset)
MEF	mouse embryonic fibroblast
MEF-NT	reprogrammed transcriptome of MEF by <i>Xenopus</i> oocyte factors
mESC	mouse embryonic stem cell
mESC-NT	reprogrammed transcriptome of mESC by <i>Xenopus</i> oocyte factors
mMyo	mouse myoblast
mMyo-NT	reprogrammed transcriptome of mMyo by <i>Xenopus</i> oocyte factors
mSox2	mouse SRY-box 2
Oocyte-NT	nuclear transfer to oocytes
OSKM	Oct4, Sox2, Klf4 and Myc (a combination of transcription factors, used to induce pluripotency)
SLO	Streptolysin O
QPCR	Quantitative polymerase chain reaction
SCNR	somatic cell nuclear reprogramming
SCNT	somatic cell nuclear transfer
TF	transcription factor
<i>xklf2</i>	<i>Xenopus</i> kruppel-like factor 4
<i>xpou60</i>	<i>Xenopus</i> POU class 5 homeobox 3, gene 3 (aka <i>Xenopus pou5f3.3</i>)
<i>xsox2</i>	<i>Xenopus</i> SRY-box 2
YF	Yamanaka factor

Table of Contents

Preface.....	I
Summary.....	II
Acknowledgement.....	IV
Abbreviations and explanations.....	VI
Table of Contents.....	1
 Chapter 1 Introduction.....	 7
1.1 Reprogramming of somatic cell nuclei by three SCNR approaches and the phenotypic assessment.....	9
1.1.1 SCNR to totipotency.....	10
1.1.2 SCNR to pluripotency.....	12
1.1.3 SCNR to other differentiated cell types.....	15
1.2 The resistance of cells to SCNR.....	17
1.2.1 DNA methylation and SCNR resistance.....	18
1.2.2 Histones and SCNR resistance.....	21
1.2.3 Chromatin modifiers/remodelers and SCNR resistance	23
1.3 Strategies that lead to successful SCNR.....	25
1.3.1 The relationship among chromatin accessibility, TF binding and downstream gene regulation.....	26
1.3.2 The effect of the combination of TFs on chromatin accessibility	29
1.3.3 Pioneer transcription factors and inaccessible regions of chromatin	31
1.4 The Yamanaka factors in ESCs and <i>Xenopus</i> eggs/oocytes contribute to maintain and induce pluripotency.....	33
1.4.1 Yamanaka factors in pluripotent cells maintain pluripotent chromatin conformation	34
1.4.2 The redundancy of Yamanaka factors in maintenance and induction of pluripotency.....	36
1.4.3 Yamanaka factor homologs in <i>Xenopus</i> and other species during early development.....	38
1.4.4 The interspecies regulation and conservation of transcription factors in pluripotency from an evo-devo view	40
 Chapter 2 Materials and Methods.....	 42
2.1 Introduction	42
2.2 Overexpression of TF proteins in <i>Xenopus</i> oocytes.....	43
2.2.1 Plasmid construction	43
2.2.2 <i>In vitro</i> transcription.....	44
2.2.3 Western blotting and oocyte dissection	44
2.3 Nuclear transfer into <i>Xenopus laevis</i> oocytes	45
2.3.1 <i>Xenopus</i> oocyte preparation	45
2.3.2 Cell culture	46
2.3.3 Cell permeabilization.....	46
2.3.4 Nuclear transfer.....	47
2.4 Measurements of transcriptional reprogramming by oocytes and effects of overexpression of transcription factors.....	48
2.4.1 RNA extraction.....	48
2.4.2 Reverse transcription	49
2.4.3 QPCR.....	49
2.4.4 RNA immunoprecipitation	50
2.4.5 RNA-seq.....	51

Chapter 3 The effect of mammalian and <i>Xenopus</i> Yamanaka factors on somatic cell nuclear reprogramming by <i>Xenopus</i> oocytes	52
3.1 Introduction	52
3.1.1 Background	52
3.1.2 Experimental design.....	53
3.2 Mammalian Yamanaka factors can be produced in <i>Xenopus</i> oocytes after mRNA injection and regulate pluripotency genes in MEFs at Day 2 after Oocyte-NT	55
3.2.1 The production and localization of mammalian YFs after mRNA injection in <i>Xenopus</i> oocytes.....	55
3.2.2 The overexpression of mammalian YFs regulates pluripotency genes in MEFs at Day 2 after Oocyte-NT	73
3.2.3 Summary	78
3.3 <i>Xenopus</i> Yamanaka factors can be produced in the <i>Xenopus</i> oocytes after mRNA injection and regulate pluripotency genes in MEFs at Day 2 after Oocyte-NT	79
3.3.1 The production and localization of <i>Xenopus</i> YFs after mRNA injection in <i>Xenopus</i> oocytes.....	80
3.3.2 The overexpression of <i>Xenopus</i> YFs regulates pluripotency genes in MEFs at Day 2 after Oocyte-NT	88
3.3.3 Summary	92
3.4 Conclusions	93
Chapter 4 Time-point observation demonstrates gene regulation in mammalian nuclei by oocyte factors and overexpression of <i>xklf2</i>-HA ..	96
4.1 Introduction	96
4.1.1 Background	96
4.1.2 Experimental design.....	97
4.2 The overexpression of <i>xklf2</i>-HA regulates some tested pluripotency genes in <i>sixi</i>FM MEFs and may be an oocyte factor during SCNR by <i>Xenopus</i> oocytes	100
4.2.1 Oocyte factors up-regulate mOct4, mSox2, mKlf4, mSall4, mUtf1 and mEsrrb in <i>sixi</i> FM MEFs from Day 0 to Day 3 after Oocyte-NT	100
4.2.2 <i>xklf2</i> -HA overexpression augments the up-regulation of mOct4, mSox2, mKlf4 and mSall4 by oocyte factors by oocyte factors at Day 1, 2 and 3 after Oocyte-NT	101
4.2.3 The fluctuating augmentation caused by <i>xklf2</i> -HA overexpression explains the stochastic nature of oocyte	102
4.2.4 The up-regulation of pluripotency genes by oocyte factors happens most strongly within one day after Oocyte-NT and the one-day regulation is sufficient to represent the long-term regulation from Day 0 to Day 3	103
4.2.5 Up-regulation of mOct4, mSox2, mKlf4, mSall4 and mUtf1 by <i>xklf2</i> -HA overexpression starts at different time points after Oocyte-NT	104
4.2.6 <i>xklf2</i> -HA overexpression up-regulates some selected pluripotency genes as oocyte factors and it may be one of the effective oocyte factors	105
4.2.7 Summary	109
4.3 The overexpression of <i>xklf2</i>-HA facilitates the expression of pluripotency genes in MEFs at the beginning of SCNR by <i>Xenopus</i> oocytes	110
4.3.1 Up-regulation of pluripotency genes in TcR2 MEFs by oocyte factors and <i>xklf2</i> -HA overexpression is similar to the up-regulation of genes in <i>sixi</i> FM MEFs	110
4.3.2 Early and late response of pluripotency genes to <i>xklf2</i> -HA overexpression at Day 0 and Day 2 after Oocyte-NT	111

4.3.3 mOct4 is a SCNR resistant gene in MEFs and is up-regulated by xklf2-HA overexpression immediately at Day 0 after Oocyte-NT with the highest fold change among all test pluripotency genes.....	112
4.3.4 Summary.....	116
4.4 The similar and different response of pluripotency genes in MEFs and mMyoblasts to oocyte factors and to xklf2-HA overexpression	117
4.4.1 Similar response of tested pluripotency genes in MEFs and mMyos to oocyte factors and xklf2-HA overexpression except for the response of mUtf1 in mMyos to oocyte factors	117
4.4.2 mOct4 and mSox2 in mMyos are up-regulated by oocyte factors and xklf2-HA overexpression more strongly than mOct4 and mSox2 in MEFs	119
4.4.3 The nil or low expressed genes in donor cells are more susceptible to the up-regulation induced by oocyte factors and xklf2-HA overexpression	121
4.4.4 Summary.....	125
4.5 Conclusions	126
Chapter 5 Oocytes reprogram gene expression of transcriptomes of mESCs, MEFs and mMyos to an oocyte-steady state except for cell-type specific genes	131
5.1 Introduction	131
5.1.1 Background	131
5.1.2 Experimental design.....	133
5.2 <i>Xenopus</i> oocytes reprogram MEFs to a steady state within two days after Oocyte-NT	136
5.2.1 The Day 1 and Day 2 samples of MEF-NT are similar and the expression of only 0.5% of analyzed genes is changed by <i>Xenopus</i> oocytes from Day 1 to Day 2 after Oocyte-NT	136
5.2.2 The expression of 66% newly synthesized transcripts in MEF-NT remains constant from Day 1 to Day 2 after Oocyte-NT and some pluripotency genes are up-regulated by oocyte factors within one day after Oocyte-NT	139
5.2.3 The batch effects affect the speed of SCNR by <i>Xenopus</i> oocytes but different batches of oocytes still can reprogram transcriptomes of MEFs to a highly similar oocyte-steady state at Day 2 after Oocyte-NT	144
5.2.4 Summary	148
5.3 <i>Xenopus</i> oocytes reprogram the transcriptomes of mESCs, MEFs and mMyos to be almost identical at Day 2 after Oocyte-NT	151
5.3.1 mESC-NT, MEF-NT and mMyo-NT are not grouped by the difference of cell types or different origins of oocyte batches	152
5.3.2 There are less than 40 Cell-type DE genes differentially expressed between reprogrammed transcriptomes of mESCs, MEFs and mMyos at Day 2 after Oocyte-NT.....	153
5.3.3 At Day 2 after Oocyte-NT, more than 12000 genes in mESC-NT, MEF-NT and mMyo-NT are reprogrammed by oocyte factors to the same level and, in addition to those genes, 1000-6000 genes are cell-type specific in the donor cells	158
5.3.4 Summary	163
5.4 Transcriptional reprogramming of mESCs, MEFs and mMyos by <i>Xenopus</i> oocytes	165
5.4.1 Large-scale change of transcriptomes of mESCs, MEFs and mMyos before and after Oocyte-NT	165
5.4.2 Some genes of mESCs, MEFs and mMyos are silent during SCNR by <i>Xenopus</i> oocytes.....	172
5.4.3 Summary	172

5.5 Distinguish between reprogrammable genes and cell-type specific genes among mESC-NT, MEF-NT and mMyo-NT and the extent to which oocytes activate silent genes	174
5.5.1 Reprogrammed transcriptomes of mESCs, MEFs and mMyos contain reprogrammable genes and cell-type specific genes.....	174
5.5.2 Reprogrammable genes possess the ability to change chromatin structures and regulate genes as the new functions of oocyte-determined cell identity.....	181
5.5.3 Summary	181
5.6 Conclusions	183
Chapter 6 <i>Xenopus</i> klf2-HA overexpression facilitates regulation of target genes in mESCs, MEFs and mMyos during SCNR by oocytes ..	184
6.1 Background	184
6.1.1 Introduction	184
6.1.2 Experimental design.....	186
6.2 Overexpression of xklf2-HA reprograms MEF transcriptome within 1 day after Oocyte-NT	189
6.2.1 xklf2-HA acts mostly as a transcriptional activator and xklf2-HA overexpression reprogram MEF transcriptome within 1 day after Oocyte-NT ..	189
6.2.2 Transcriptional reprogramming of transplanted cells by xklf2-HA overexpression happens mostly within 24 hours after Oocyte-NT	194
6.2.3 Summary	198
6.3 Overexpression of xklf2-HA facilitates the regulation of gene expression in mESCs, MEFs and mMyos during SCNR by oocytes.....	199
6.3.1 Overexpression of xklf2-HA regulates more than 1000 genes in mESCs, MEFs and mMyos at Day 2 after Oocyte-NT	200
6.3.2 xklf2-HA overexpression facilitates regulation of downstream genes after Oocyte-NT in MEF-NT	205
6.3.3 Summary	207
6.4 Overexpression of xklf2-HA regulates xklf2-DE genes to the xklf2-oocyte level in mESC-NT, MEF-NT and mMyo-NT at Day 2 after Oocyte-NT	208
6.4.1 xklf2-HA overexpression regulates different sets of xklf2-DE genes in mESC-NT, MFE-NT and mMyo-NT at Day 2 after Oocyte-NT while xklf2-DE genes are either up-regulated or down-regulated by xklf2-HA overexpression	208
6.4.2 xklf2-HA overexpression up-regulates expression of xklf2-DE genes to xklf2-oocyte level.....	213
6.4.3 Summary	218
6.5 Up-regulated xklf2-DE genes by xklf2-HA overexpression during SCNR by oocytes are responsible for developmental processes and signaling pathways regulating pluripotency of stem cells.....	219
6.5.1 The majority of xklf2-DE genes have mKlf2 binding motifs in their promoters and xklf2-DE genes are signaling pathways regulating pluripotency of stem cells are enriched for up-regulated xklf2-DE genes	220
6.5.2 Summary	223
6.6 xklf2-HA overexpression activates SCNR resistant genes in mESCs, MEFs and mMyos.....	225
6.6.1 42~46% of xklf2-DE genes are reprogrammable genes and are regulated by both maternal factors and xklf2-HA overexpression	226
6.6.2 5~11% of cell-type specific genes resist to SCNR by oocytes but are activated by xklf2-HA overexpression within 2 days after Oocyte-NT	227
6.6.3 Reprogrammed transcriptomes contain up-regulated and down-regulated xklf2-DE genes but only up-regulated xklf2-DE genes contribute to activate resistant genes.....	228

6.6.4 xklf2-HA overexpression activates and regulates genes responsible for signaling pathways regulating pluripotency of stem cells	233
6.6.5 Summary	237
6.7 Conclusions	239
Chapter 7 The effect of the pioneer transcription factor, mFoxa1-HA, on somatic cell nuclear reprogramming by <i>Xenopus</i> oocytes	241
7.1 Introduction	241
7.1.1 Background	241
7.1.2 Experimental design.....	242
7.2 mFoxa1-HA accumulates in the GV dose-dependently and SUMOylation of mFoxa1-HA is observed in the <i>Xenopus</i> oocytes	244
7.2.1 mFoxa1-HA proteins accumulate in the GV and increase time- and dose-dependently in different cellular compartments	245
7.2.2 SUMOylation is observed in mFoxa1-HA protein detection	251
7.2.3 Summary	255
7.3 The overexpression of mFoxa1-HA up-regulates mSox2 only in MEFs during SCNR by <i>Xenopus</i> oocytes.....	256
7.3.1 The overexpression of mFoxa1-HA up-regulates mSox2 among tested pluripotency genes	256
7.3.2 The overexpression of mFoxa1-HA up-regulates tested pluripotency genes with variable relative expression in each biological replicate	258
7.3.3 Summary	260
7.4 The regulation of neurogenic genes by mFoxa1 is different between Oocyte-NT and neurogenesis of mDA neurons/progenitors	261
7.4.1 The overexpression of mFoxa1-HA up-regulates mDmrtb1, mFoxb1 and mOtx2 in sixiFM MEF up to 31-fold at Day 2 after Oocyte-NT	264
7.4.2 Oocyte factors up-regulate mOtx2, mFerd3l, mKcnip3, mOtx1, mGli1 in sixiFM MEF up to 313-fold from Day 0 to Day 2 after Oocyte-NT.....	266
7.4.3 mFoxa1-HA overexpression up-regulates mDmrtb1, mFoxb1 and mOtx2 in sixiFM MEF with variable relative expression due to SCNR resistance	270
7.4.4 Summary	273
7.5 Conclusions	274
Chapter 8 Discussions	276
8.1 Introduction	276
8.2 Gene reprogramming by maternal factors of <i>Xenopus</i> oocytes and by TF overexpression	278
8.3 Cell-hybrid experiments	300
Chapter 9 References	303
Chapter 10 Appendices	314
10.1 Appendix I: Plasmid construction.....	314
10.1.1 Construction of pTOPO-mFoxa1 plasmid from pBabe-puro-FoxA1 plasmid and pTOPO vector via Gateway cloning system	314
10.1.2 Fixation of the missense mutation of mFoxa1 coding sequence on pTOPO-mFoxa1 plasmid	317
10.1.3 Construction of pCS2-mFoxa1-3HA plasmid from pTOPO-mFoxa1 plasmid and pCS2-3HA-attR vector	319
10.2 Appendix II: Primer list for neurogenic genes	320
10.3 Appendix III: Bioinformatics analysis	321
10.3.1 Hybrid genome generation.....	321
10.3.2 Processing of fastq files	321
10.3.3 Reproducibility analysis.....	321
10.3.4 Differential expression (DE) analysis.	322

10.4 Appendix IV: Preliminary data validation for reprogrammed transcriptomes of mESCs, MEFs and mMyos by <i>Xenopus</i> oocytes.....	323
10.5 Appendix V: Gene ontology and KEGG pathway analysis of silenced genes in mESCs, MEFs and mMyos after Oocyte-NT	327
10.6 Appendix VI: Gene ontology and KEGG pathway analysis of reprogrammable genes in mESCs, MEFs and mMyos after Oocyte-NT	337
10.7 Appendix VII: Preliminary data validation for reprogrammed transcriptomes of mESCs, MEFs and mMyos in the absence and presence of xklf2-HA overexpression in the Oocyte-NT system.....	344
10.8 Appendix VIII: The involvement of xklf2-HA downstream genes in gene regulation during SCNR by oocytes	349
10.9 Appendix IX: xklf2-HA overexpression activates genes at nil/low expression level and represses genes at high expression level during SCNR by <i>Xenopus</i> oocytes	353
10.10 Appendix X: xklf2-HA overexpression down-regulates xklf2-DE genes to xklf2-oocyte level at Day 2 after Oocyte-NT	361
10.11 Appendix XI: Gene ontology and KEGG pathway analysis for up-regulated and down-regulated xklf2-DE genes	365
10.12 Appendix XII: 6.6 xklf2-HA overexpression activates resistant genes and up-regulates the expression of genes to xklf2-oocyte level	381

Chapter 1 Introduction

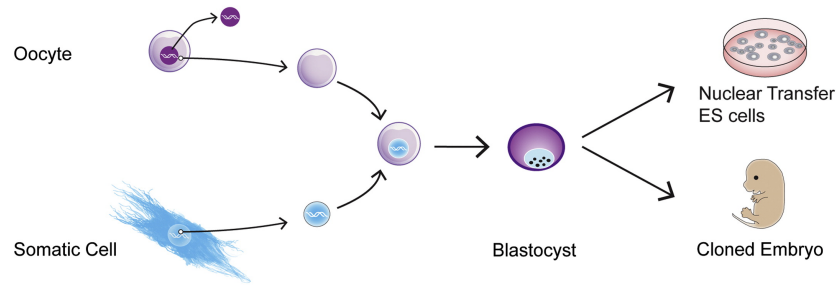
Somatic cell nuclear reprogramming (SCNR) is the process by which terminally differentiated cells revert to a stem cell state or turn into other cell-types¹. Three approaches can be utilized to achieve SCNR, including somatic cell nuclear transfer, cell fusion and overexpression of transcription factors (Figure 1.1)^{1,2}. Through these SCNR approaches, researchers can investigate the mechanistic details behind SCNR and have a better idea how to apply SCNR techniques to regenerative medicine, drug discoveries and other applications^{3,4}.

Mechanistic studies have uncovered the roles of reprogramming factors in epigenetic and transcriptional changes during SCNR¹. The involvement of the appropriate reprogramming factors at the different phases and the removal of inadequate memories in somatic cell nuclei lead to successful SCNR¹. However, the efficiency of SCNR is usually low and the mechanisms requires clarification².

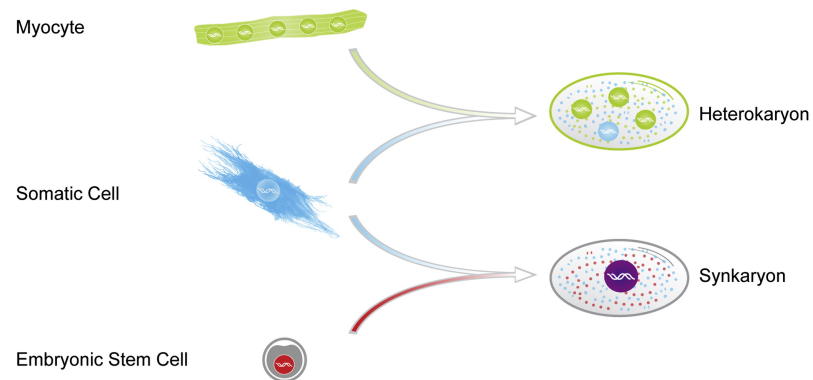
In order to further understand the mechanistic details of SCNR, my project compares the difference in transcriptional reprogramming between native maternal factors and overexpression of transcription factors via somatic cell nuclear transfer in non-dividing *Xenopus* oocytes. From this comparison, I would like to ask if transcription factors help to induce totipotency and how they promote SCNR by *Xenopus* oocytes?

Figure 1.1

A Somatic Cell Nuclear Transfer



B Cell Fusion



C Transcription Factor Transduction

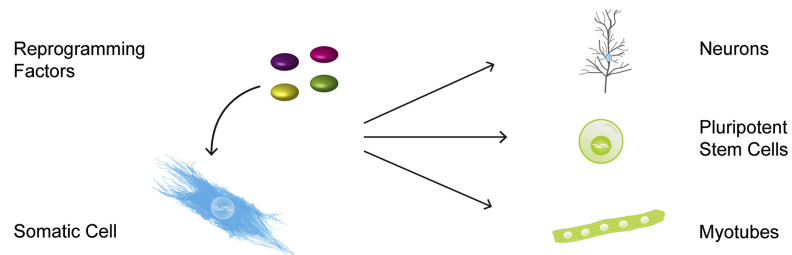


Figure 1.1 Experimental systems for studying nuclear reprogramming².

(A) Somatic cell nuclear transfer

(B) Cell fusion

(C) Transcription factor transduction

1.1 Reprogramming of somatic cell nuclei by three SCNR approaches and the phenotypic assessment

For SCNR approaches, reprogramming factors are provided to induce and support the SCNR process. These reprogramming factors might be modulators of chromatin conformation, such as DNA methyltransferases/demethylases, canonical histones and histone variants, chromatin modifiers/remodelers, and are capable of changing the accessibility of chromatin^{1,5}. To reprogram somatic cells to other cell identities, transcriptional regulators, namely transcription factors, cofactors and non-coding RNAs, are also needed and are able to change the characteristics of cells^{1,4}.

In these SCNR studies, it is sometimes arguable that the plasticity of donor cells are uncertain while cells with higher plasticity, rather than terminally differentiated somatic cells, are used for SCNR and these cells are more susceptible to reprogramming factors¹. To clearly understand SCNR, this section introduces studies that utilized three SCNR approaches and includes the information about the donor cell types, the reprogrammed cells and the ways to evaluate SCNR.

1.1.1 SCNR to totipotency

Totipotency is the capability of a cell to give rise to cells in both embryonic and extra-embryonic lineages⁶. Therefore, the criterion used to assess the totipotent state of a cell is to confirm if this cell/zygote can generate an entire organism⁶. In nature, totipotent zygotes can be formed from oocytes/eggs and sperm through fertilization. Artificially, cloning an adult from somatic nuclei can also be achieved via nuclear transfer to eggs and it has been achieved in many species^{7,8}.

While applying nuclear transfer to eggs, components in eggs empower somatic nuclei regain totipotency and develop to adult animals⁹. In the early 1950s, Briggs and King performed nuclear transfer to enucleated eggs from blastula and gastrula cells and the resulting cells yield tadpoles of *Rana Pipiens* successfully^{10,11}. Later, Gurdon and colleagues managed to produce sexually mature frogs from embryonic and somatic cell nuclei in *Xenopus laevis*¹²⁻¹⁵. After decades of effort, the first mammalian clone was successfully derived from adult differentiated cells of mammary glands and was named as Dolly the sheep in 1997¹⁶. In 2018, the first non-human primate was cloned from fetal fibroblasts through modifying contents in donor eggs^{17,18}.

Notably, forming a totipotent zygote through fertilization is usually highly efficient except for the failure of fertilization caused by the deterioration or defects of gametes. However, forming a totipotent cell through somatic cell nuclear transfer is of low efficiency due to the resistance of donor cell nuclei to reprogramming triggered by maternal factors of eggs. Since both routes allow

a nucleus to become totipotent, the different efficiency between fertilization and somatic cell nuclear transfer into eggs suggests the resistance of somatic cell nuclei to the maternal reprogramming factors in eggs.

To study the maternal effect on SCNR, extracts of oocytes/eggs are useful to investigate the immediate response. Somatic nuclei are exposed to maternal factors of oocyte/egg extracts. Also maternal factors from oocyte/egg extracts are shown to epigenetically change the somatic cell nuclei and change the nuclear structure to improve the SCNR^{19,20}. Intriguingly, cytoplasm of GV-oocytes (Prophase I) and MII-oocytes (Metaphase II) affects the somatic cell nuclei in different ways. For example, the activity of Pol I and Pol II polymerases is maintained when incubating the somatic cell nuclei in the cytoplasm of GV-oocytes but the activity is abolished in the cytoplasm of MII-oocytes¹⁹.

In addition, the cytoplasm of GV-oocytes may be more potent to induce pluripotency than MII-oocytes since the cytoplasm of GV-oocytes can activate and maintain the expression of NANOG in fibroblasts after 7 days of extract treatment while cytoplasm of MII-oocytes can only activate but not maintain the NANOG expression²¹. Furthermore, it has been shown that the exposure of fibroblasts to a GV-oocyte cytoplasmic extract can epigenetically improve somatic cell nuclear reprogramming because it can increase the efficiency of somatic cell nuclear transfer into MII-oocytes²².

1.1.2 SCNR to pluripotency

After fertilization or nuclear transfer to eggs, the totipotent zygotes/cells divide and, after rounds of cell division, some of these cells become pluripotent²³. Pluripotent cells are able to derive cells that form three germ layers and, in embryos, pluripotent cells can be found at characterized anatomical positions^{23,24}. During the transition from totipotency to pluripotency, it has been found that the mobility of core histones decreases as development proceeds and pluripotent cells retain higher mobility of histones than the differentiated trophectoderm²⁵.

There are several types of cultured pluripotent stem cells, including embryonic stem cells (ESCs), epiblast-derived stem cells (EpiSCs), embryonic germ cells and embryonal carcinoma cells²⁶. Several findings point out that these pluripotent stem cells are interconvertible by modulating cell-intrinsic and exogenous factors²⁶. One *in vivo* way to examine if a cell is pluripotent is to transplant it to the blastocyst and confirm its pluripotency by the formation of a chimera²⁴. The other *in vivo* way is to implant the cell subcutaneously into kidney a capsule and see if it can form a teratoma²⁴.

ESCs are one type of pluripotent stem cells and are derived from the inner cell mass of blastocysts²⁴. In mouse, pluripotent states of ESCs can be divided to naïve and primed pluripotency in pre-implantation blastocysts and post-implantation embryos respectively²⁷. Interestingly, human ESCs have only primed pluripotency and naïve pluripotency of human ESCs can be

derived experimentally by resetting transcription factor control circuitry or modified culture conditions²⁸⁻³⁰.

Although ESCs can be a good resource for making all the functional cells and can be applied in medicine, they can only be used in research due to ethical issues. Therefore, induced pluripotent stem cells were derived by forced expression of combinatorial sets of transcription factors, mainly Yamanaka factors, in various differentiated cells³¹. Likewise, iPSCs were proven to generate the cells of all three germ layers via *in vivo* assays and have the benefit of producing personalized iPSCs for regenerative medicine and drug discovery³.

In 2006, Yamanaka and Takahashi found as few as four transcription factors, Oct3/4, Sox2, c-Myc and Klf4, can reprogram embryonic and adult fibroblasts to pluripotent state³². Later, it has been shown that Nanog-expressed iPSCs are germline-competent and iPSCs can be derived from terminally differentiated B lymphocytes accompanied by the interruption of transcriptional state maintaining B lymphocyte identity^{33,34}. For the equivalence between ESCs and iPSCs, although it remains controversial, ESCs and iPSCs were reported to be transcriptionally and epigenetically identical when they have the same genetic origin^{26,35}.

In addition to induced pluripotency, cell fusion can also be utilized to study pluripotency²³. By fusing somatic cells with pluripotent stem cells, the pluripotency genes in somatic cells are activated and the reprogrammed

fusion hybrid cells can differentiate to three germ layers *in vitro* and *in vivo*³⁶. In the absence of cell division, Do and Schöler fused either karyoplasts of ESCs or cytoplasts of ESCs with neurosphere cells and found the Oct4 gene was only activated by karyoplasts³⁷. Additionally, Nanog was shown to promote transfer of pluripotency after fusion of ESCs and differentiated cell nuclei while elevated Nanog improves the yield of hybrid cell colonies with ESC characteristics³⁸.

The other pluripotent stem cells, cell extracts of embryonal carcinoma cells, were shown to up-regulate the expression of OCT4 and NANOG genes and caused the demethylation of OCT4 promoter and enhancer³⁹. Also, the acetylation of H3K9 and demethylation of H3K9 and H3K27 occur during this transcriptional activation³⁹.

For the activation of silenced genes by cell fusion, a biphasic remodeling was proposed⁴⁰. The early phase of the activation of silenced genes is triggered by transcription activators and repressors independent of cell division and this happens when fusing somatic cells with either another type of somatic cell or ESCs⁴⁰. The late phase of activation of silenced genes involves a DNA methylation changes and chromatin remodeling of silenced genes which relies on cell division and this only happens when fusing somatic cells with ESCs⁴⁰.

1.1.3 SCNR to other differentiated cell types

The third type of SCNR reprograms somatic cells to cell-types of the same or related cell lineages, providing a resource for regenerative medicine⁴. Two routes can be involved in this type of SCNR⁴. One is dedifferentiation and the other is transdifferentiation⁴. Dedifferentiation and transdifferentiation happen naturally in many animals, such as in heart regeneration in zebrafish, limb regeneration in urodele amphibians and Schwann cell proliferation in mammals⁴. Resulting cells could be the progenitors of the reprogrammed somatic cells or other types of differentiated functional cells⁴. Additionally, upon tissue injury, adult cells can dedifferentiate and gain plasticity to differentiate into other cell types⁴¹.

Mechanistically, cell fusion can be used to study transdifferentiation. In 1983, Blau and colleagues fused human amniocytes with mouse muscle cells and made a stable heterokaryon that did not undergo cell division⁴². They also showed that muscle genes in human amniocytes could be activated by the cytoplasm of mouse muscle cells when the nuclei of these two cell types remained distinct in the heterokaryon⁴². Later, the same group proved that muscle gene activation in non-muscle cells by muscle cells is independent of DNA synthesis and the kinetics of gene activation are related to the plasticity of non-muscle cells and the germ layer difference^{43,44}.

Forced dedifferentiation and transdifferentiation can be achieved by the overexpression of transcription factors⁴. Ways to evaluate the successful dedifferentiation and transdifferentiation of reprogrammed somatic cells are to

examine the expression of cell/tissue specific genes, global expression profiles and to perform cell/tissue specific functional assays^{45,46}. For example, fibroblasts can directly transdifferentiate into cardiomyocytes by the combinatorial overexpression of Gata4, Mef2 and Tbx5⁴⁵. Another way to transdifferentiate cells to other cell types is to dedifferentiate the original cell identity followed by differentiation condition. Take lens regeneration in newts as an example, Margariti and colleagues partially dedifferentiated fibroblasts *via* an induced pluripotency approach and then differentiated the partially dedifferentiated fibroblasts into endothelial cells by VEGF treatment⁴⁶.

1.2 The resistance of cells to SCNR

During development, cells become more specialized and ultimately differentiate to be functional as a variety of cell types. The genome-wide study showed that developmental specification is accompanied by chromatin restriction progressively⁴⁷. Furthermore, exposure of collected human tissues and stem cells to serum triggers a distinct epigenome transition, involving *de novo* establishment of domains with features of constitutive heterochromatin⁴⁷.

From totipotent zygotes/cells, pluripotent stem cells, multipotent stem cells, unipotent progenitors to terminally differentiated cells, chromatin changes its conformation to be more restricted to certain cell-types along the developmental processes and influence the cell fate decision and reprogramming⁴⁸. However, the maintenance and establishment of chromatin states for specialized cell identities become the barriers for SCNR^{2,49}.

Insufficient removal of epigenetic barriers for somatic cell nuclear reprogramming causes the low yield of cloned animals and abnormalities in developmental embryos and adults⁵. Mechanistically, chromatin dynamics during iPSC reprogramming were studied and it has been pointed out that complete epigenetic reprogramming is needed to prime for the transcriptional reprogramming⁵⁰⁻⁵². In this section, epigenetic barriers involved in SCNR will be introduced.

1.2.1 DNA methylation and SCNR resistance

The DNA methylome of cells changes dynamically during the life cycle of animals and is responsible for genome stability and gene expression of specified cell-types⁵³. Throughout development, global DNA demethylation occurs after fertilization and during the early phase of gametogenesis⁵³. Afterwards, the DNA methylome builds up again for cell lineage specification and X-inactivation⁵⁴.

In terms of the relationship between DNA methylation and cell plasticity, it has been demonstrated that DNA demethylation is related to totipotency and pluripotency during early development and gametogenesis⁵⁴. After fertilization, DNA demethylation occurs globally, including in the imprinted regions of genes, while the totipotent zygote is formed⁵⁴. Then, the totipotent zygote divides into pluripotent cells of inner cell mass. The pluripotent cells keep on dividing and become pluripotent EpiSCs after X-inactivation (epiblast-derived stem cells) and pluripotent PGCs (primordial germ cells). During gametogenesis, PGCs then experience another DNA demethylation, genomic imprinting and become totipotency-competent germ cells again⁵³⁻⁵⁵.

In addition, the stability of the DNA methylome has been shown to be indispensable for self-renewal of multipotent stem cells and the dynamics of DNA methylation is the determinant of tissue homeostasis and an abnormal change of DNA methylome may cause cancer⁵⁶. However, despite the fact that X-inactivation is one of the crucial DNA methylation events during development, and occurs during the transition from naïve pluripotency to

primed pluripotency, its role remains unclear. To help explain this, Chen and colleagues showed that the X-inactivation was not tightly synchronized with differentiation although sex-related gene expression varies widely across these stages⁵⁷.

Since the DNA methylation restricts plasticity of cells and responsible for stability of the epigenome, it suggests that changes in DNA methylome are essential for SCNR⁵⁸. A systematic cell-fusion study shows that DNA methylation functions as the memory of silenced genes in fibroblasts in the presence of transcriptional activators from other cell types⁵⁹.

By fusion of non-muscle HeLa cells and muscle cells, muscle genes in HeLa nuclei were only activated under the treatment of the demethylation agent, 5-azacytidine, and genes were activated sequentially and interdependently⁶⁰. It has also been shown that 5-azacytidine can increase the efficiency of induced pluripotency⁶¹. Furthermore, DNA demethylase, AID, was shown to remove DNA methylation at promoters of pluripotency genes and is required for active DNA demethylation and initiation of nuclear reprogramming towards pluripotency in human somatic cells⁶².

On the other hand, DNA methylation was shown to correlate with chromatin structure and to affect chromatin accessibility through methyl-CpG-binding domain (MBD) proteins^{58,63}. It has been demonstrated that the MBD proteins Mbd3 can form a repressor complex with NuRD (nucleosome remodeling and

deacetylase) and enhance the efficiency of induced pluripotency to 100% within 7 days⁶³.

Interestingly, although DNA methylation increases during development, DNA methylation was also gained gradually along the process of SCNR in the induced pluripotency model⁶⁴. After SCNR, a loss of DNA methylation is achieved instead when somatic cells become ESC-like cells and DNA methylation was shown to be a switch of pluripotency genes⁶⁴. CpG-rich promoters of some pluripotency genes retain low methylation with strong engagement of histone marks and CpG-poor promoters of other pluripotency genes are safeguarded by methylation at an early stage of reprogramming⁶⁴.

1.2.2 Histones and SCNR resistance

It has long been known that histones wrap DNA and form nucleosomes as the basic unit of chromatin structure. To further compact DNA, nucleosomes interact with other regulatory proteins to form higher-order chromatin structures. There are two kinds of histones involved in the construction of chromatin structure - replication-dependent and replication-independent histones. These histones can change the chromatin structure by deposition and removal machineries and contribute to normal development, lineage commitment of stem cells and SCNR^{48,65}.

During development, the mobility of histones is correlated with the plasticity of cells. It has been shown that the mobility of core histones, H2A, H3.1, and H3.2, is unusually high in totipotent cells in two-cell embryos, and mobility decreases while totipotent cells develop into pluripotent cells²⁵. During SCNR, histone variant macroH2A was shown to mark the differentiated cell states during mouse embryogenesis and removal of macroH2A increases the efficiency of induced pluripotency⁶⁶. Interestingly, chromatin assembly by replication-dependent histones can repress basal transcription but can not repress the function of certain transactivators, such as Gal4-VP16⁶⁷.

In addition, variable histone modifications mark different chromatin regions and are crucial to transcriptional regulation and other functions⁶⁸. For example, trimethylation of H3 lysine 4 (H3K4me3), H3K4me3 is a mark specifying the promoter of active genes. The breadth and location of it are linked to cell identities and broad H3K4me3 domains help to identify essential genes in

different cell-types⁶⁹. The same study also pointed out that breadth of H3K4me3 correlates with the consistent expression of transcription factors, rather than with the expression level of transcription factors when comparing various cell-types⁶⁹.

For the repressive marks, trimethylation of H3 lysine 9 (H3K9me3), although it was traditionally associated with noncoding regions of the genome and studies have shown that it is also crucial to repress lineage-inappropriate genes and therefore impede SCNR⁷⁰. It has been shown that ectopic expressed H3K9me3 demethylase, Kdm4d, could activate SCNR resistant genes and improves SCNR efficiency⁷¹. Additionally, SCNR efficiency can be improved by the depletion of H3K9 methyltransferases⁷¹.

Furthermore, combining sonication-resistant heterochromatin with H3K9me3 marked regions, these regions could more accurately predict the resistant genes to activation during direct cell type conversion than the genes marked by histone modification alone⁷². A recent breakthrough in SCNR shows that cloned non-human primate can be achieved by injection of H3K9me3 demethylase Kdm4d mRNA and treatment with histone deacetylase inhibitor trichostatin A at one-cell stage following somatic cell nuclear transfer¹⁷.

1.2.3 Chromatin modifiers/remodelers and SCNR resistance

Global chromatin remodeling during differentiation leads to a progressive transition from a fairly open chromatin configuration to a more compact state and chromatin modifiers and remodelers play important roles in these developmental transitions⁷³. Conversely, these chromatin states with varied plasticity are also critical to SCNR⁴⁸.

It has been reported that chromatin modifiers can be barriers or facilitators of SCNR⁷⁴. When inducing pluripotency, inhibition of core components of polycomb repressive complex 1 and 2 results in both repression and enhancement of SCNR efficiency⁷⁴. More specifically, inhibition of H3K27 methyltransferase EZH2 reduced SCNR efficiency and suppression of SUV39H1, YY1 and DOT1L (H3K79 methyltransferase) enhanced reprogramming⁷⁴.

Moreover, histone chaperone CAF-1 was proven to safeguard somatic cell identity during induced pluripotency⁷⁵. Optimal modulation of CAF-1 and transcription factor levels increased efficiency of induced pluripotency and facilitated iPSC formation in as little as 4 days⁷⁵. Mechanistically, CAF-1 suppression increases the accessibility of enhancers early during SCNR, followed by a decrease in somatic heterochromatin domains, an increase of Sox2 binding to pluripotent genes and activation of associated genes⁷⁵.

Interestingly, the pluripotent cells can be reprogrammed into a totipotent state by downregulating the replication-dependent chromatin assembly⁷⁶. Ishiuchi

and colleagues showed that reprogramming from ESCs to totipotent 2C-like cells can be achieved by knockdown of the chromatin-assembly factor, CAF-1 and the atypical histone deposition can contribute to the reprogramming towards totipotency⁷⁶.

1.3 Strategies that lead to successful SCNR

In previous section, it has discussed about the epigenetic barriers during SCNR. One genome-wide analysis for barriers in iPSC generation identified not only epigenetic barriers but also other putative barriers, such as genes involved in transcription, ubiquitination, dephosphorylation, vesicular transport and cell adhesion and inhibition of these barriers led to improved SCNR efficiency^{77,78}. Another genome-wide study revealed factors needed at transition steps of induced pluripotency⁷⁹. This study showed that some critical genes are associated with cell signaling pathways before cells acquire pluripotency and a specific gene set is important to mature iPSC formation⁷⁹.

For the establishment and maintenance of cell identities, the epigenetic states and transcriptional regulation are both crucial. During SCNR, transcription factors are mediators of reprogramming and the reestablishment of transcription regulator bindings on chromatin during cell division aids the transcriptional reprogramming⁸⁰. Additionally, due to massive barriers during induced pluripotency, Yamanaka factors (OCT4, SOX2, KLF4 and c-MYC: OSKM) have been proposed as pioneer factors that can bind to target sequences reside in inaccessible chromatin⁸¹. In this section, the mechanisms relates to the transcriptional regulators during SCNR will be discussed.

1.3.1 The relationship among chromatin accessibility, TF binding and downstream gene regulation

In *Drosophila*, a systematic study has shown that five principal chromatin types have been characterized by 53 selected chromatin components and DNA-binding factors bind to their cognate motifs differentially guided by these characterized chromatin types⁶⁸. During *Drosophila* embryo development, it has been shown that the accessible regions of chromatin, DNase I hypersensitive sites, are associated with spatial-temporal expression pattern of linked genes⁸². Additionally, the level of occupancy of transcription factors in early *Drosophila* embryos on their target genes is more correlated with the degree of chromatin accessibility *in vivo* than the *in vitro* affinity measurements using purified protein and naked DNA^{83,84}. Moreover, in highly accessible chromatin, transcription factors at a sufficiently high concentration can occupy their recognition sites without the physical cooperative aid from other proteins and so lead to widespread and overlapped binding patterns by transcription factors⁸⁴.

In early development of *Xenopus*, histone synthesis is not coordinated with DNA synthesis⁸⁵. In *Xenopus* oocytes, replication-dependent histones are synthesized without DNA replication and the replication-dependent histones can repress basal gene expression and the ratio of histones to DNA represents the accessibility of DNA to transcription regulators^{67,85}. After fertilization, histones are synthesized at a rate far more than immediate requirements for synthesized DNA and there is only a low amount of maternal gene expression and no zygotic gene expression due to the constraints of

chromatin accessibility to the transcription factors before mid-blastula transition^{67,85}. It has been shown in *Xenopus* that the titration of H3/H4 histones with an excessive amount of DNA can cause early activation of zygotic genes before mid-blastula transition⁸⁶. Furthermore, the competition between unbound histones and transcription factors has been shown to cause early onset of zygotic gene expression before MBT in zebrafish⁸⁷.

In the same context, an increase or decrease in the amount of transcription factors can also affect the competition between histones and transcription factors and lead to the activation of silenced genes by transcription factors. During the self-renewal of pluripotent ESCs with a fixed landscape of chromatin structure, the amount of transcription factors controls the cell fate decision delicately^{88,89}. A less than twofold increase of Oct-3/4 can cause the differentiation of mouse ESCs to primitive endoderm and mesoderm⁸⁸; repression of Oct-3/4 can cause the loss of pluripotency and dedifferentiation of mouse ESCs to trophectoderm⁸⁸. Additionally, a minor increase (less than twofold) of Sox2 in mouse ESCs leads to the differentiation to a wide range of cell-types, including neuroectoderm, mesoderm, and trophectoderm but not endoderm⁸⁹.

Interestingly, in the process of induced pluripotency, different levels of expression of KLF4 pause the partially reprogrammed iPSCs at successive stages and it may relate to the onset of different pluripotency genes following the increase of the expression of KLF4⁹⁰. In their study, Nishimura and colleagues analysed the effect of decreasing expression levels of OCT4,

SOX2, KLF4 and c-MYC and found that the expression level of KLF4 reproducibly derives homogenous populations of partially reprogrammed iPSCs⁹⁰. The different expression levels of KLF4 pause partially reprogrammed iPSCs at distinct intermediate stages and up-regulation of KLF4 leads cells to resume reprogramming and increases expression levels of ESC markers, Cdh1, Fgf4 and Rex1⁹⁰.

1.3.2 The effect of the combination of TFs on chromatin accessibility

The binding of transcription factors to their recognition sites is correlated with chromatin accessibility during development and winning the competition with the unbound histones can lead to the reactivation of silenced genes^{82-84,87}. In terms of the selection of a combination of transcription factors, this is unrelated to the cell-types and species since the combination of *Oct4*, *Sox2*, *Klf4* and *c-Myc* (OSKM) gives rise to mouse iPSCs from several cell-types and the homologs of OSKM gives rise to the iPSCs in many species³¹.

Furthermore, different combinations of transcription factors can also induce pluripotency although the majority of the combinations are still based on the combination of OSKM³¹. For example, the combination of OCT4, SOX2, NANOG and LIN28 can also be used to induce human iPSCs^{31,91}. Interestingly, the OSKM based combination can be interchangeable with homologs of other species, such as the utilization of mouse OSKM to induce *Xenopus* iPSCs *in vivo*⁹².

Since the binding of transcription factors largely relies on the landscape of chromatin accessibility, the increase of chromatin accessibility is essential for induction of pluripotency by OSKM at the beginning. With the aid of cell division, induced pluripotency is hypothesized to reverse this developmentally-imposed repression and it is believed to immensely change the chromatin structure of terminally differentiated somatic cells to a pluripotent state by combinations of few transcription factors⁸⁰. It has been demonstrated that the inhibition of the p53/p21 pathway or overexpression of

Lin28 could increase the cell division rate and lead to accelerated kinetics of iPSC formation⁹³. Interestingly, Nanog overexpression can also accelerate induced pluripotency reprogramming in a predominantly cell-division-rate-independent manner⁹³.

Additionally, the combination of Oct4, Sox2 and Klf4 (OSK) without c-Myc has been proven to reset the epigenomes of somatic cells by upregulating the expression of several chromatin remodelers and chromatin modifiers⁹⁴. Furthermore, the cooperative binding of OSK on enhancers of somatic genes early in the reprogramming process immediately initiates the redistribution of somatic TFs on enhancers of somatic genes and results in the repression of them by recruiting Hdac1⁹⁵. Some of the enhancers of pluripotency genes are bound by OSK in early reprogramming but most of them are selected later in reprogramming by OS and other pluripotency genes⁹⁵.

Altogether, the combination of OSKM to induce pluripotency from various somatic cells is versatile in many species³¹. The combination of OSKM can lead to the similar regulation of downstream genes via the cooperative interaction with other transcription regulators from the beginning to the completion of induced pluripotency and can reconstruct the chromatin structures to be ESC-like as a result³⁵.

1.3.3 Pioneer transcription factors and inaccessible regions of chromatin

During development, chromatin accessibility defines the binding of TFs and the regulation of downstream genes by TFs. However, the combination of OSKM can reverse this process through mechanisms involving the change of landscape of chromatin structure⁹³⁻⁹⁵. It is understandable that the combination of OSKM may lead to the bindings to regulatory elements of their target genes. Conversely, it may be ideal that one or more of OSKM are pioneer transcription factors and can open closed chromatin, bind to target sequences and recruit other transcription factors to regulate necessary downstream genes for inducing pluripotency.

In development, evidences have emerged for the roles of pioneer transcription factors in ESCs and early stages of development⁹⁶. Take Foxa1 as an example, Foxa1 is a liver-specifying gene and expression initiates in the gut endoderm at E7/E8 of mouse embryos⁹⁷. Foxa1 virtually regulates all the liver-specific genes and the binding of Foxa1 to inactive genes indicates its pioneer capability⁹⁷. It has been shown that the early stages of liver development depend on Foxa1 and Foxa2 *in vivo* and the primary cells from Foxa1/Foxa2-deficient endoderm fail to initiate expression of liver-specific genes under the stimulation of exogenous FGF2 *in vitro*⁹⁸. Therefore, Foxa1 is competent to initiate expression of liver-specific genes.

It has been reported that the high nucleosome-binding affinity of Foxa1/2 contributes to the slower nuclear mobility of Foxa1/2 when compared with other transcription factors⁹⁹. In addition to that, the slower nuclear mobility of

Foxa1/2 is caused by its non-specific binding due to the resemblance between DNA binding domain of Foxa proteins and linker histone⁹⁹. However, it has been proven that Foxa1 may not always be the first transcription factor entering the closed chromatin and the bindings of steroid receptors are sometimes the first to enhance the binding of Foxa1¹⁰⁰.

Although it is controversial, pioneer factor activity may still be controlled by transcription factors, which can contribute to the SCNR. One example is the transdifferentiation of terminally differentiated fibroblasts or hepatocytes to induced neurons by the combination of Ascl1, Brn2 and Myt1l^{101,102}. Ascl1 is shown to be a pioneer factor among these¹⁰³.

Furthermore, OSK, as potential pioneer factors, have been shown to promote induced pluripotency by binding to enhancers of silent pluripotency genes marked with H3K9me3, and these H3K9me3 marked regions prevent the initial binding of OSKM and impede the efficiency of reprogramming¹⁰⁴. Moreover, the partial nucleosome-binding affinity of OSK has been reported to be evidence of OSK as pioneer factors¹⁰⁵. By contrast, the change of epigenome during induced pluripotency is shown to be a downstream effect *via* the upregulation of several chromatin modifiers and remodellers⁹⁴. Moreover, enhancers of silent pluripotency genes can only be bound by OSK cooperatively while there is no binding to enhancers by these pluripotency genes by overexpression of a single TF⁹⁵.

1.4 The Yamanaka factors in ESCs and *Xenopus* eggs/oocytes contribute to maintain and induce pluripotency

During induced pluripotency, it is evident that the overexpression of OSKM leads to ESC-like chromatin structures regardless of the original cell-types, although the mechanisms for the reconstruction of chromatin towards a pluripotency state by the combination of OSKM are elusive^{31,35}. Therefore, it implies that diverse chromatin structures of various cell-types would be destined to uniform ESC-like chromatin structure by the combinatorial regulation of OSKM with the help of unknown mechanisms¹⁰⁶. Then, the ESC-like chromatin structures would be susceptible to certain extracellular stimuli and differentiate into all kinds of somatic cells.

Interestingly, nuclear transfer into eggs (Egg-NT) follows the same conceptual transitions except that a totipotent chromatin structure is set as a result and the SCNR determinants for Egg-NT are more complicated than the combination of Yamanaka factors for induced pluripotency. For epigenetic reprogramming, the chromatin structural change by Egg-NT has been discussed in previous sections and in reviews^{5,26}. In this section, I will introduce the roles of Yamanaka factors in maintaining and inducing pluripotency.

1.4.1 Yamanaka factors in pluripotent cells maintain pluripotent chromatin conformation

Pluripotency is a state that exists transiently in the early embryo and can be recapitulated *in vitro* by deriving embryonic stem cells or by induction of pluripotency in somatic cells¹⁰⁷. The state of pluripotency is stabilized by a gene regulatory network of transcription factors and the epigenome of pluripotent cells is proposed to be under the control of TF network¹⁰⁸. By integrating external signals and exerting control over the decision between self-renewal and differentiation, the imbalance of TF network would lead to the exit from pluripotency¹⁰⁷.

It has been reported that cell fusion with pluripotent cells would epigenetically reprogram the somatic cells, including DNA demethylation and reactivation of an X chromosome^{62,109}. This means the factors within the pluripotent cells execute this epigenetic reprogramming process and maintain the self-renewal of pluripotent cells at the same time. For the maintenance of ESCs, it has been shown that Oct3/4 and Sox2 cooperatively activate Lefty1 by binding to its enhancers, mediated by the binding of Klf4 on its promoter¹¹⁰. Since Oct3/4 and Sox2 are the core factors that supports the self-renewal of ESCs, the precise level of them is critical to govern the distinct fates of ESCs towards differentiation, dedifferentiation or self-renewal^{88,89,111}. A small increase in the level of Sox2 would lead to the inhibition of its own expression and down-regulate some Sox2:Oct-3/4 downstream genes possibly via feedback and feedforward gene regulation control^{89,111}.

The binding of Yamanaka factors determines the continuity of the self-renewal circuit and the chromatin structure of pluripotent stem cells can be maintained by a core regulatory network in defined culture medium¹¹². The TF network and signaling pathways defines the core regulation of a cell and explains the maintenance of pluripotent chromatin structure and pluripotent transcriptomes¹¹². In mice, ESCs cultured in 2i/LIF are naïve-pluripotent and EpiSCs cultured in FGF/activinA medium are primed-pluripotent¹¹². Under 2i/LIF conditions, Oct4 primarily induces Klf2 expression, compared with Klf4 expression, and LIF/Stat3 selectively enhances Klf4 expression¹¹³. The LIF dependence is reduced by overexpression of either Klf2 or Klf4 and the expression of either of these can reinstate naïve pluripotency from primed-pluripotent EpiSCs¹¹³. Additionally, Klf2 sustains ground state pluripotency via the inhibition of Mek/Erk pathways while both prodifferentiation Mek/Erk and Gsk3/Tcf3 are inhibited under 2i/LIF condition¹¹⁴. Moreover, the Mek/Erk/Klf2 pathway cooperates with the Gsk3/Tcf3/Esrrb pathway in mediating naïve pluripotency¹¹⁴.

1.4.2 The redundancy of Yamanaka factors in maintenance and induction of pluripotency

Since the DNA binding domains of TFs belong to the same families and some members of a TF family have an identical or similar canonical binding motif, the functions of TFs would not totally depend on recognition of binding sites but also be context-dependent in maintenance and induction of pluripotency, early development and cancer¹¹⁵⁻¹¹⁷. The context-dependent functions of TFs in the same family are sometimes redundant while some members of the same TF family are interchangeable in the same context¹¹⁵⁻¹¹⁷. For Yamanaka factors, some family members can compensate them for sustaining self-renewal of pluripotent stem cells and inducing pluripotency.

Klf2, 4 and 5 have been shown to redundantly bind and regulate crucial pluripotency genes in mouse ESCs and to sustain self-renewal of ESCs¹¹⁸. Therefore, the spontaneous differentiation of mouse ESCs can only be achieved under triple siRNA knockdown of Klf2, 4 and 5 altogether while the triple siRNA knockdown leads to the loss of self-renewal of mouse ESCs¹¹⁸.

Nevertheless, it has also been reported that Klf2, 4 and 5 regulate different sets of genes in mouse ESCs and therefore play different roles in self-renewal of mouse ESCs^{119,120}. Moreover, inactivation of Klf5 by zinc finger nuclease has been shown to downregulate expression of pluripotency genes and to attenuate colony formation in mouse ESCs¹²¹. It has further been indicated that Klf4 and Klf5 differentially bind and repress endoderm and mesoderm markers in mouse ESCs, respectively, and knockdown of either of these

genes leads to the enhancement of differentiation towards endoderm or mesoderm in early embryo development¹²².

During induced pluripotency and self-renewal of pluripotent stem cells, SOX1, SOX3 and re-engineered SOX7 and SOX17 have been reported to be functionally redundant with SOX2¹¹⁷. Interestingly, Sox7 and Sox17 can be converted to be reprogramming factors by a single amino-acid alteration within its DNA binding domain, HMG. Sox7EK and Sox17EK with a single amino-acid mutation from glutamate (E) to lysine (K) can replace Sox2 and interact with Oct4 to induce pluripotency while Sox4EK, Sox5EK, Sox6EK, Sox8EK, Sox9EK, Sox11EK, Sox12EK, Sox13EK and Sox18EK cannot¹²³. Moreover, the C-terminal transactivation domain of Sox7 and Sox17 enhances the potency of Sox2 to induce pluripotency and confers weak reprogramming potential of mutated Sox4EK and Sox18EK¹²³.

1.4.3 Yamanaka factor homologs in *Xenopus* and other species during early development

Since we are using the *Xenopus* system to study SCNR in our lab, it is useful to understand the roles of Yamanaka factor homologs in early development of different species and evaluate the conservation of these factors in our system.

The pluripotent gene, POU5F1 (OCT4), is a key regulator of stem cell fate and homologues of POU5F1 present throughout vertebrates¹¹⁶. In *Xenopus*, transcripts of members of Pou family are very abundant during early embryogenesis and absent from most somatic tissues¹²⁴. During early embryo development, xpou60, xpou25 and xpou91 are highly related to mammalian Oct-3/4 and sequentially expressed¹²⁴. Transcripts of xpou60 are maternally expressed in *Xenopus* oocytes and the xpou60 proteins are present in oocytes and in early embryos until gastrulation^{124,125}. Transcripts of xpou25 are present at a low level in oocytes and early embryos and are upregulated greatly in early gastrulation but xpou25 proteins are detected during gastrulation, not in oocytes and early embryos¹²⁴. Transcripts of xpou91 are induced after midblastula transition and reach the highest level in late gastrulation¹²⁴.

Since xpou60 is expressed in oocytes but not expressed in somatic cells and the oocyte extract can activate xpou60 in somatic cells through the Oct-Sox binding motif, it suggests the expression of xpou60 in oocytes is possibly regulated by a positive feedback loop¹²⁵. Furthermore, the expression of xpou25, xpou60 and xpou91 decrease during late gastrulation and early

neurulation due to the repression by *sall4* for promoting posterior neural fates¹²⁶.

In the mouse, *Oct4* is restricted to inner cell mass and plays a key role in lineage segregation of trophectoderm and inner cell mass¹²⁷. However, in the pig, *Oct4* is not restricted to inner cell mass cells and *Sox2* becomes a pluripotency marker instead since *Sox2* restricts to the inner cell mass¹²⁷. Therefore, transcription factors related to lineage specification may be variable among species and the corresponding TF networks may differ as a result although the phenotypic outcomes are the same.

During *Xenopus* embryogenesis, members of Klf families, such as *xklf2*, *xklf4*, *xklf5*, *xklf6*, *xklf7*, *xklf8*, *xklf11*, *xklf15* and *xklf17*, are involved in germ layer formation and body axis patterning^{128,129}. It has been shown that these Klf family members are expressed in early *Xenopus* embryos with different spatial-temporal patterns^{128,129}. Interestingly, each Klf member regulates distinct sets of genes that are essential for germ layer specification and body axis formation and sometimes there is an overlapping regulation of these downstream genes among different Klf members^{128,129}. This explains the redundancy of some transcription factors of the same families while some of their downstream genes are overlapping but some other downstream genes are differently induced at the molecular level.

1.4.4 The interspecies regulation and conservation of transcription factors in pluripotency from an evo-devo view

Although conserved Yamanaka factors can be found not only in human and mouse but also in *Xenopus* and other species, the interspecies difference in transcriptional regulation of these factors is unclear. Interestingly, one study compared ChIP-seq among five vertebrates to determine the genome-wide occupancy of two TFs, C/EBP alpha and HNF4 alpha¹³⁰. It showed that most binding is species-specific and aligned binding events present in all five species are rare although the DNA binding preferences are highly conserved for each TF¹³⁰. Binding divergence between species can be explained by sequence changes to the bound motif and the interspecies transcriptional regulation is different evolutionarily¹³⁰.

For the homologues of pluripotent factor POU5F1, they are functionally conserved to regulate early embryonic potency and differentiation and exist throughout vertebrates. Mouse Pou5f1 (Oct3/4) was firstly identified among class V members of POU family and is a central regulator of ESC pluripotency and induced pluripotency¹¹⁶. Notably, some vertebrate lineages have both POU5F1 and POU5F3 orthologues whereas others have either only POU5F1 or only POU5F3¹¹⁶. In *Xenopus*, three *POU5F3* genes but no *POU5F1* genes are present. These three *POU5F3* genes are *pou5f3.1* (*xpou91*), *pou5f3.2* (*xpou25*) and *pou5f3.3* (*xpou60*), and only *pou5f3.1* is expressed in primordial germ cells and possesses OCT4-like activity in iPSC reprogramming and ESC self-renewal¹¹⁶.

The *Pou2* and *Oct4* genes evolved from a *POU class V* gene ancestor. Some *Pou2* and *Oct4* homologues of different vertebrates, such as medaka and axolotl, were shown to induce pluripotency in mouse and human fibroblasts¹³¹. These results indicate that induction of pluripotency is not restricted to mammals but exists in the *Oct4/Pou2* ancestral vertebrate¹³¹.

It has been reported that PouV proteins of *Xenopus* and the axolotl, namely xpou25, xpou60, xpou91 and amOct4, can bind to an octamer-binding motif and support self-renewal of mouse Oct4-depleted ESCs¹³². Mouse Oct4 and axolotl Oct4 can rescue the xpouV-deleted phenotype of *Xenopus* embryos¹³². In addition, transcriptional activation of mouse Oct4 and *Xenopus* xpou91 by fusion with the activation domain, VP60, supports maintenance and induction of pluripotency and alleviates the dependence of LIF in ESC maintenance medium¹³³.

Interestingly, induced pluripotency can be achieved by direct delivery of mouse OSK into *Xenopus* muscle fibers⁹². *Xenopus* pluripotency genes can be activated by mouse OSK and xventx2, which functionally resembles Nanog, is induced consequently⁹². Furthermore, these pNanog-GFP positive cells can differentiate into derivatives of three embryonic lineages *in vitro* and into neurons and muscles *in vivo*⁹². While *Xenopus* do not have a *NANOG* gene, axolotl *Nanog* activity was proven to support the ground state pluripotency of mouse ESCs and demonstrates that the mechanisms governing pluripotency are conserved from urodele amphibians to mammals¹³⁴.

Chapter 2 Materials and Methods

2.1 Introduction

In our lab, we use nuclear transfer to oocytes (Oocyte-NT) to study the possible mechanisms involved in somatic cell nuclear reprogramming.

In my thesis, I investigated the similarity and difference between nuclear transfer to oocytes and overexpression of transcription factors (TFs) through time course observations. By directly overexpressing single transcription factor in oocytes before Oocyte-NT, the transcriptional reprogramming induced by oocyte factors and the effect of TF overexpression on gene regulation during this process is compared.

At first, I tested six Yamanaka homologs and one pioneer TF and selected *xklf2*-HA to perform the genome-wide evaluation via RNA-seq plus BrUTP pull-down. In addition, I used three cell types to perform Oocyte-NT and tried to find the nuclear reprogramming resistant genes in different cell types.

2.2 Overexpression of TF proteins in *Xenopus* oocytes

Thanks to Jullien J, Garrett NJ and previous lab member, Pasque V, I received 6 plasmids of Yamanaka factors, including hOCT4-HA, mSox2-HA, hKLF4, xpou60-HA, xsox2-HA and xklf2-HA, in the laboratory repository. These plasmids can be used directly for *in vitro* transcription. For the pioneer transcription factor mFoxa1, the original plasmid pBabe-puro-FoxA1 was a gift from Prof Kenneth S. Zaret. The wild-type mFoxa1 on pBabe-puro-FoxA1 contains a missense mutation possibly due to long-term cultivation so I mutated it back (Appendix I, page 318) referred to the published sequence (NM_008259.3) and added an HA tag to it through Gateway cloning system, as instructed by Jullien J and as referred to in manufacturers' protocols.

2.2.1 Plasmid construction

The coding sequence of mFoxa1 in pBabe-puro-FoxA1 plasmid was first amplified by high-fidelity DNA polymerase (Phusion High-Fidelity DNA polymerase, 2U/μl, Cat No M0530, New England Biolabs). Since the mFoxa1 coding sequence of pBabe-puro-FoxA1 plasmid was spotted with a missense mutation from Ser-12 to Asn-12, the mutation of plasmid was firstly mutated back to the original Ser-12. To add a T7 promoter and HA tags to the N-terminus of mFoxa1, the mFoxa1 coding sequence on pTOPO-mFoxa1 plasmid was recombined with pCS2-3HA-attR vector via LR clonase (Gateway LR clonase II enzyme mix, Thermo Fisher Scientific).

2.2.2 *In vitro* transcription

Each constructed plasmid that contains the coding sequence of a transcription factor and a T7 promoter was used as starting materials for *in vitro* transcription. First, constructed plasmids were linearized by restriction enzymes and treated with proteinase K (800 units/ml, NEB). The linearized plasmids were then purified by phenol/chloroform extraction and ethanol precipitation. The resulting products were then used for producing TF mRNA with the RNA cap analog (m⁷G(5')ppp(5')G RNA cap structure analog, NEB) via *in vitro* transcription (MEGAscript T7 transcription kit, Thermo Fisher Scientific). The mRNA yield was then purified by spin-column chromatography (Illustra Microspin G-50 column, GE Healthcare), phenol/chloroform extraction and ethanol precipitation. The mRNA pellet then dissolved in DNase/RNase free H₂O, aliquotted, snap-frozen on dry ice and stored in -70°C freezer.

2.2.3 Western blotting and oocyte dissection

Having *in vitro* synthesized TF mRNA, the TF mRNA was injected into cytosols of *Xenopus* oocytes to produce TF proteins and TF protein production and localization can be evaluated via Western blotting. First, different doses of TF mRNA were injected into oocytes and collected at different time points after incubating the mRNA-injected oocytes at 18°C. Then, the oocytes were dissected in mineral oil and the fractions of cytosols and GVs were collected separately to confirm newly-made TF proteins are transported into GVs. Western blotting was then run to probe the newly-made TF proteins in different fractions and LI-COR Odyssey® CLx Imaging System was utilized to quantify the relative amount of newly-made TF proteins.

2.3 Nuclear transfer into *Xenopus laevis* oocytes

Nuclear transfer into *Xenopus laevis* oocytes (Oocyte-NT) starts with preparing *Xenopus* oocytes, which were individualized and detached from layers of follicular cells by liberase treatment. Before Oocyte-NT, the oocytes of experimental groups were injected with TF mRNA and incubated at 18°C for 24 hours. The donor cells used for Oocyte-NT were then permeabilized with streptolysin O (SLO) and the holes of permeabilized cells allow oocyte factors and newly-made TF proteins to enter the cells immediately after Oocyte-NT. The experimental details refer to the paper of our previous lab member, Halley-Stott RP¹³⁵.

2.3.1 *Xenopus* oocyte preparation

In all the experiments, the stage V/VI oocytes were used. Firstly, the ovaries were taken from the adult female *Xenopus Laevis*. The oocytes connected with stroma tissue were torn apart as strings by forceps and put into 50ml conical tubes. The strings of oocytes were then treated with liberase (Roche, 28U/ml in H₂O) for 2 hours with gentle agitation (15 rpm). Then the liberase was washed away from the oocytes with 1X MBS and the stage V/VI oocytes with diameter range of 1 to 1.2 mm were selected for the following experiments. The selected oocytes were placed in petri dish in 1X MBS at 16°C and the follicular cell layer would detach in an overnight incubation. The oocytes were then stored in 1X MBS with 0.1% FBS at 16°C and ready for manipulation.

2.3.2 Cell culture

I have used four cell lines for nuclear transfer. B10 CHD4 is a mouse embryonic stem cell (mESC) line, TcR2 and sixiFM are mouse embryonic fibroblast (MEF) cell lines and C2C12 is mouse myoblast (mMyo) cell line.

B10 CHD4, requires medium renewal everyday and medium contains GMEM, 20% fetal bovine serum, leukemia inhibitory factor, non-essential amino acid, β -mercaptoethanol, sodium pyruvate, penicillin, streptomycin and fungizone.

The flasks or dishes used for B10 CHD4 require gelatin coating before seeding. For MEF and mMyo cell lines, the medium needs to be renewed every two to three days and it contains DMEM, 10% fetal bovine serum, penicillin, streptomycin and fungizone. Cells were subcultured or frozen when they reached subconfluence and incubated at 37°C with 5% CO₂.

2.3.3 Cell permeabilization

Cells were cultured to semi- to sub-confluence. Cells were washed twice with PBS and detached with trypsin for 5 to 10 minutes at 37°C. Trypsin was neutralized with DMEM and cells were individualized by pipetting. Cells were centrifuged at 500 rpm for four minutes, the supernatant was discarded and the cells were resuspended with PBS. Cells were centrifuged at 2000 rpm for one minute and PBS was replaced with SuNaSp solution. Cells were centrifuged at 2000 rpm for one minute and the supernatant was discarded. 20ul SLO (20,000 units/ml in PBS, containing 0.01% BSA and 5mM DTT, Sigma-Aldrich, S5265) was added plus 100ul SuNaSp solution for 3-6x10⁶ cells¹³⁵. Cells were then resuspended by pipetting and permeabilized at 37°C

in a waterbath for one minute. Cells were incubated on ice and some cells were taken to check permeabilization efficiency (95~99%) under a microscope by trypan blue staining. SLO reaction was stopped by adding SuNaSp BSA. Cells were centrifuged at 2000 rpm for one minute, the supernatant was discarded and then resuspended with SuNaSp BSA solution. Cells in SuNaSp BSA solution were then aliquotted, snap-frozen on dry ice and stored in a -70 °C freezer.

2.3.4 Nuclear transfer

Pull the needles with suitable size for nuclear transfer. Mix cell suspension with plasmid encodes cytoplasmic membrane GFP. Inject 9.2nl cell suspension (300~500 cells) with 5pg plasmid of cytoplasmic membrane GFP into germinal vesicle of each oocyte. During Oocyte-NT, the permeabilized cells were sucked into micropipettes via Drummond Nanoinject injector and injected into GV of oocytes. Select GFP positive oocytes 24 hours after nuclear transfer. Collect samples at needed time points, snap-freeze samples on dry ice and store samples in -70°C freezer.

2.4 Measurements of transcriptional reprogramming by oocytes and effects of overexpression of transcription factors

Transcriptional changes after nuclear transfer are measured by two methods. One is quantitative PCR (QPCR) and the other is RNA-seq plus BrUTP pull-down. Through time-point observations, QPCR shows the change of relative amount of transcripts in the donor cells and after nuclear transfer and RNA-seq plus BrUTP shows the change of relative amount of newly-synthesized transcripts after nuclear transfer. The effect of overexpression of transcription factors can be evaluated by comparing the expression level of genes between nuclear transfer samples in the presence and absence of overexpressed transcription factors by both methods. The protocols are adapted and experimental details can refer to papers of our lab members and Nature protocols¹³⁵⁻¹³⁷.

2.4.1 RNA extraction

Collect Oocyte-NT samples as groups of ten for RNA extraction. Qiagen RNeasy Mini Kit is used for RNA extraction and procedures are modified for our purpose. Briefly, lyse Oocyte-NT samples with 900ul RLT buffer and vortex for 4 minutes at 4°C. Add 900ul 70% ethanol and transfer mixture to RNeasy spin column. Centrifuge at 10,000 rpm for 30 seconds and discard flow-through. Add 350ul RW1 buffer and centrifuge at 10,000 rpm for 30 seconds. Discard flow-through and add 80ul DNase I incubation mix. Incubate at room temperature for 15 minutes, add 350ul RW1 buffer and centrifuge at 10,000 rpm for 30 seconds. Discard flow-through and add 500ul RPE buffer and

centrifuge at 10,000 rpm for 2 minutes. Place RNeasy spin column in new 2ml collection tubes and centrifuge at full speed for 1 minute. Place RNeasy spin column in new 1.5ml eppendorf and add 50ul RNase-free H₂O. Centrifuge at 10,000 rpm for 1 minute and measure RNA concentration by Nanodrop. Snap-freeze RNA extracts on dry ice and store RNA extract in -70°C freezer.

2.4.2 Reverse transcription

SuperScript III Reverse Transcriptase Kit is used for reverse transcription. Briefly, mix 5ug RNA extract with 1ul gene-specific primer (100uM, Sigma-Aldrich) and 1ul dNTP mix (10mM) and add RNase-free H₂O to total 13ul. Heat mixture at 65°C for 5 minutes and incubate on ice for at least 1 minute. Add 4ul 5X FS buffer, 1ul 0.1M DTT, 1ul RNase inhibitor and 1ul RNase-free H₂O into mixture. Pipette mixture gently. Incubate mixture at 55°C for 60 minutes and increase to 70°C for 15 minutes. Store cDNA in -20°C freezer.

2.4.3 QPCR

SYBR Green Ready Mix (Sigma-Aldrich) is used for QPCR. Primers for neurogenic genes are listed in the Appendix II (page 324) and primers for pluripotency genes are listed in Halley-Stott's paper¹³⁵. For most experiments, 50ng cDNA per well of a QPCR plate is used and cDNA is increased for genes of interest which are difficult to detect. mGapdh is used for normalization since the expression level of mGapdh is not changed before and after nuclear transfer and can be a indication of the number of transplanted donor cells.

2.4.4 RNA immunoprecipitation

For labeling newly synthesized RNA and to exclude carry over in donor cells, BrUTP is injected into oocytes two hours after Oocyte-NT. Oocyte-NT samples are collected and RNA is extracted as previously mentioned. For immunoprecipitating BrUTP-labelled RNA, the first step is to prepare anti-BrUTP conjugated agarose beads. Wash 100ul beads (for 5 samples) twice with 1ml Buffer I and block beads with 500ul Blocking buffer at 4°C for 1.5 hours. Centrifuge beads solution at 3000 rpm for 3.5 minutes at 4°C and remove supernatant. Make RIP buffer by mixing 1170ul Binding buffer and 30ul SUPERase•In RNase Inhibitor and distribute 200ul RIP buffer into 5 eppendorf for immunoprecipitation. Use 25-50ug RNA extract per sample and add 2.5ul SUPERase In RNase Inhibitor into each sample. Heat RNA extract at 65°C for 5 minutes, incubate on ice for at least 1 minute and spin down. Immunoprecipitate RNA with 200ul RIP buffer overnight at 4°C. Wash RNA bead mixture with Low salt buffer once, with High salt buffer twice and TET buffer once. Elute immunoprecipitated RNA with 100ul Elution buffer by incubating at room temperature for 1 minute. Centrifuge at 3000 rpm for 4 minutes and collect supernatant. Repeat the elution steps 4 times. Extract eluted RNA by phenol/chloroform extraction and ethanol precipitation. Clean up RNA extract with Qiagen RNeasy Plus Micro Kit and measure concentration of RNA extract with Qubit RNA HS Assay Kit.

2.4.5 RNA-seq

Ovation Single Cell RNA-seq System is used to prepare RNA-seq libraries from newly-synthesized RNA. 10ng of newly-synthesized RNA is used for each sample preparation. Follow the steps provided by manufacturers. cDNA reverse transcribed from newly-synthesized RNA is then obtained and amplified as RNA-seq libraries. RNA-seq libraries were validated by Tape Station and sequenced on Illumina HiSeq 2000 and 4000 for SE 50. RNA-seq results were processed by our lab member, Simeone A (Appendix III, page 325). Then I analysed results further by Excel, Matlab, Gene ontology, Venn diagram and TRNAFAC.

Chapter 3 The effect of mammalian and *Xenopus* Yamanaka factors on somatic cell nuclear reprogramming by *Xenopus* oocytes

3.1 Introduction

3.1.1 Background

Somatic cell nuclear reprogramming (SCNR) is the process by which the terminally differentiated cell nuclei are reversed to totipotent or pluripotent states. This can be achieved by transferring a somatic cell nucleus into a mature and unfertilized egg (Egg-NT)¹² or overexpressing a combination of three to four Yamanaka factors (YFs) or other transcription factors into somatic cells (induced pluripotency)³². Even now, the efficiency of SCNR via Egg-NT or induced pluripotency is low and the mechanistic details of how they work are unclear. To elucidate the mechanistic details of SCNR, especially the role of transcription factors, I combine both systems by overexpressing one of the Yamanaka factors before Oocyte-NT and evaluate how each of these affects SCNR by *Xenopus* oocytes.

In this chapter, I selected six YFs, including three mammalian YFs and three *Xenopus* YFs. The reason for choosing two sets of YFs is that Oocyte-NT is usually an interspecies interaction when the transplanted nuclei do not belong to *Xenopus laevis*. A comparison of homologs between *Xenopus* and

mammals helps to elucidate how conserved the function of an YF is in the context of SCNR.

3.1.2 Experimental design

Before performing Oocyte-NT, the production and localization of each YF in *Xenopus* oocytes were checked by Western blotting (Figure 3.1.A). Different doses of YF mRNA were injected into the cytosol of *Xenopus* oocytes and the oocyte samples were collected at different time points from 6 hours to 6 days after mRNA injection. One day and three days after mRNA injection are the main time points to check since these time points coincide with the time points when Oocyte-NT is performed and the samples are collected. The protein counterpart of each YF mRNA was made in the cytosol and transported into germinal vesicle (GV) where it executed its function. The localization of synthesized YF proteins was checked by dissecting the GV directly from the oocytes and the relative amount of each YF in the GV fraction and cytosol was compared in Western blotting analysis.

To evaluate the effect of each YF homolog on pluripotency genes during SCNR by oocytes, Oocyte-NT was performed and, 24 hours before Oocyte-NT, 9.2 ng mRNA of one YF was injected into the cytosol of *Xenopus* oocytes (Figure 3.1.B). For Oocyte-NT, sixiFM MEFs were transplanted into GVs of *Xenopus* oocytes (~300-500 cells per oocyte). Two days later, Oocyte-NT samples were collected and RNA was extracted. The expression of pluripotency genes was then examined via QPCR.

Figure 3.1

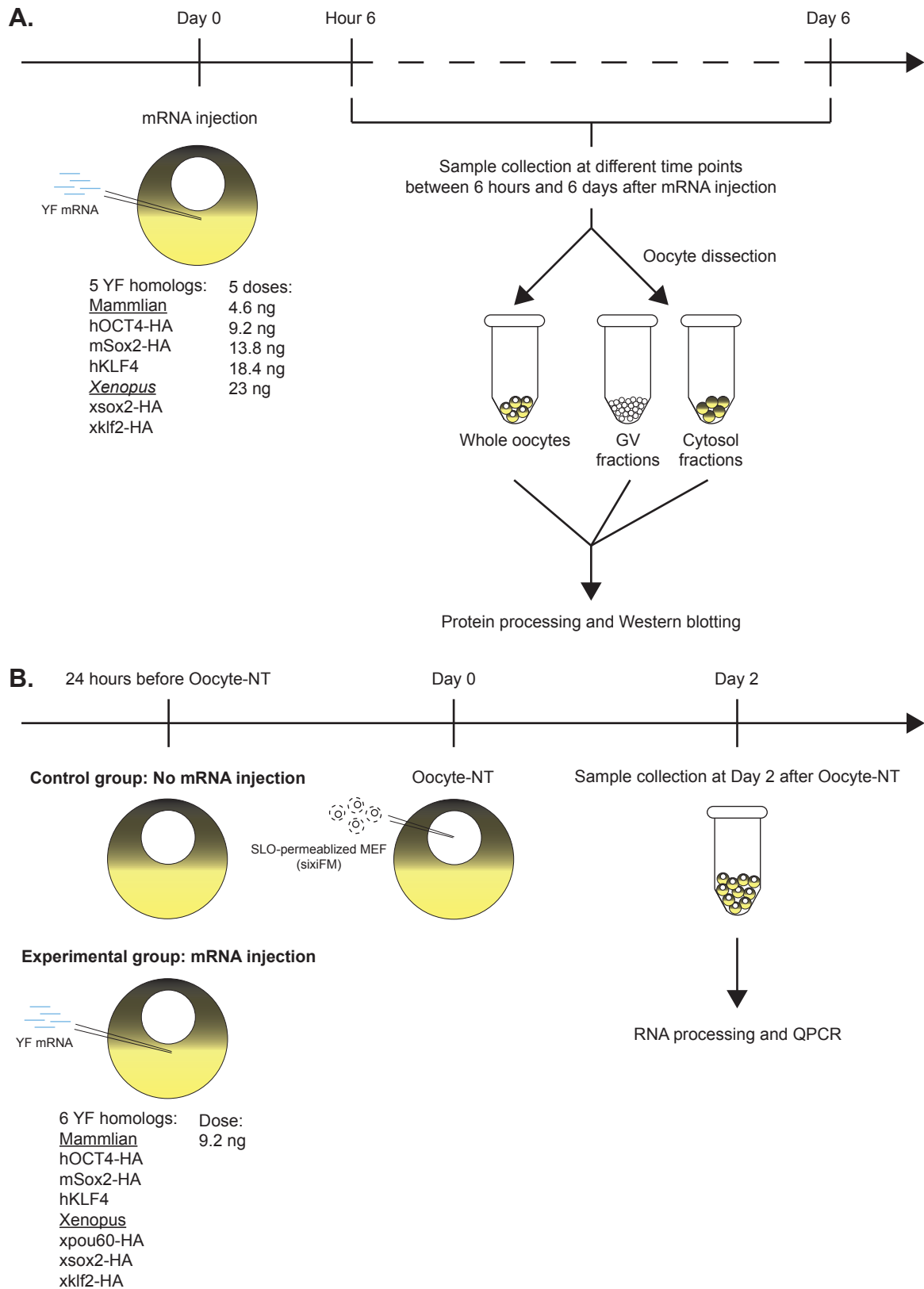


Figure 3.1 Sample preparation of Western blotting for evaluating the protein production and localization of YF homologs (A) and sample preparation of QPCR for evaluating the effect of YF homologs on pluripotency genes (B) are shown.

3.2 Mammalian Yamanaka factors can be produced in *Xenopus* oocytes after mRNA injection and regulate pluripotency genes in MEFs at Day 2 after Oocyte-NT

To evaluate the effect of mammalian YFs on pluripotency genes during SCNR by oocytes, three mammalian YFs were selected, namely hOCT4-HA, mSox2-HA and hKLF4. Because the regulation of downstream genes by transcription factors usually follows in a dose-dependent manner, the production and localization of mammalian YFs after mRNA injection were examined via Western blotting before performing Oocyte-NT (Subsection 3.2.1). After confirming that the YF proteins can be produced in the cytosol and transported into GVs, Oocyte-NT was performed (Subsection 3.2.2).

3.2.1 The production and localization of mammalian YFs after mRNA injection in *Xenopus* oocytes

To confirm if the mammalian YFs can be produced in the *Xenopus* oocytes, Western blotting was performed. Different doses of YF mRNA were injected into the cytosol of *Xenopus* oocytes at Day 0 and samples were collected at different time points after mRNA injection. The localization of newly synthesized YF proteins was evaluated by comparing the amount of YF protein in the GV fractions to the amount in the cytosol fractions or whole oocytes. The YF proteins on the Western blots were recognized by anti-HA antibody (hOCT4-HA, mSox2-HA) or by the antibody against the epitopes of the proteins (hKLF4) and anti-actin antibody was used to recognize the

internal control, actin. The amount of YF protein was quantified by LI-COR Odyssey® CLx Imaging System.

3.2.1.1 Human OCT4-HA

3.2.1.1.1 The amount of hOCT4-HA protein increases dose-dependently and decreases at Day 2 or Day 3 after mRNA injection

On the Western blot (Figure 3.2.1.A), three doses were chosen (4.6, 9.2 and 13.8 ng) for examining hOCT4-HA protein production and the samples collected at different days (Day 1, 2, 3 and 5) after mRNA injection. Each lane was loaded with a sample of 0.5 equivalent of whole oocyte. The black arrow indicates the position of hOCT4-HA protein bands and the signal intensity of hOCT4-HA protein bands is quantified and shown as numbers above the rectangles containing the quantification areas.

Comparing the samples collected at different days after injecting the oocytes with different doses of hOCT4-HA mRNA, the production of hOCT4-HA proteins is dose-dependent from 4.6 ng to 13.8 ng (Figure 3.2.1.A and B). For example, the signal intensity of Day 1 samples increases from 18700 for 4.6 ng, 38100 for 9.2 ng to 41400 for 13.8 ng. Besides, the increase of hOCT4-HA proteins in Day 1 samples reaches a plateau after the dose of hOCT4-HA mRNA at 9.2 ng since the change of signal intensity of hOCT4-HA proteins from 9.2 ng to 13.8 ng tends to level off (Figure 3.2.1.B).

However, the amount of hOCT4-HA protein decreases at Day 2 or Day 3 after mRNA injection because the signal intensity of all the doses of Day 2/Day 3 are lower than the signal intensity of corresponding doses of Day 1 samples. It is hard to judge when the amount of hOCT4-HA protein starts to decrease since the signal intensity of hOCT4-HA protein bands of Day 2 samples for each mRNA dose is less than the signal intensity of samples for corresponding doses of any other days. Therefore, it is likely that the protein transfer of Day 2 samples to the blot was not complete or the sample integrity is compromised for improper sample preparation.

The decrease of hOCT4-HA proteins at Day 2 may reduce the effectiveness of hOCT4-HA on regulating the downstream genes after Oocyte-NT performed at Day 1 after mRNA injection since the samples for Oocyte-NT are collected at Day 3 after mRNA injection. It is also possible that the decrease of hOCT4-HA proteins happens at Day 3 after mRNA injection and therefore the decrease of hOCT4-HA protein will not be a concern for Oocyte-NT.

3.2.1.1.2 hOCT4-HA proteins accumulate in the GVs dose-dependently after mRNA injection

To check the localization of hOCT4-HA proteins after they are made in the cytosol from injected mRNA, *Xenopus* oocytes were injected with three different doses of mRNA, namely 4.6, 9.2, 13.8 ng per oocyte, and the samples were collected two days after hOCT4-HA mRNA injection (Figure 3.2.1.C). Oocyte dissection is utilized to separate the GV fractions and cytosol

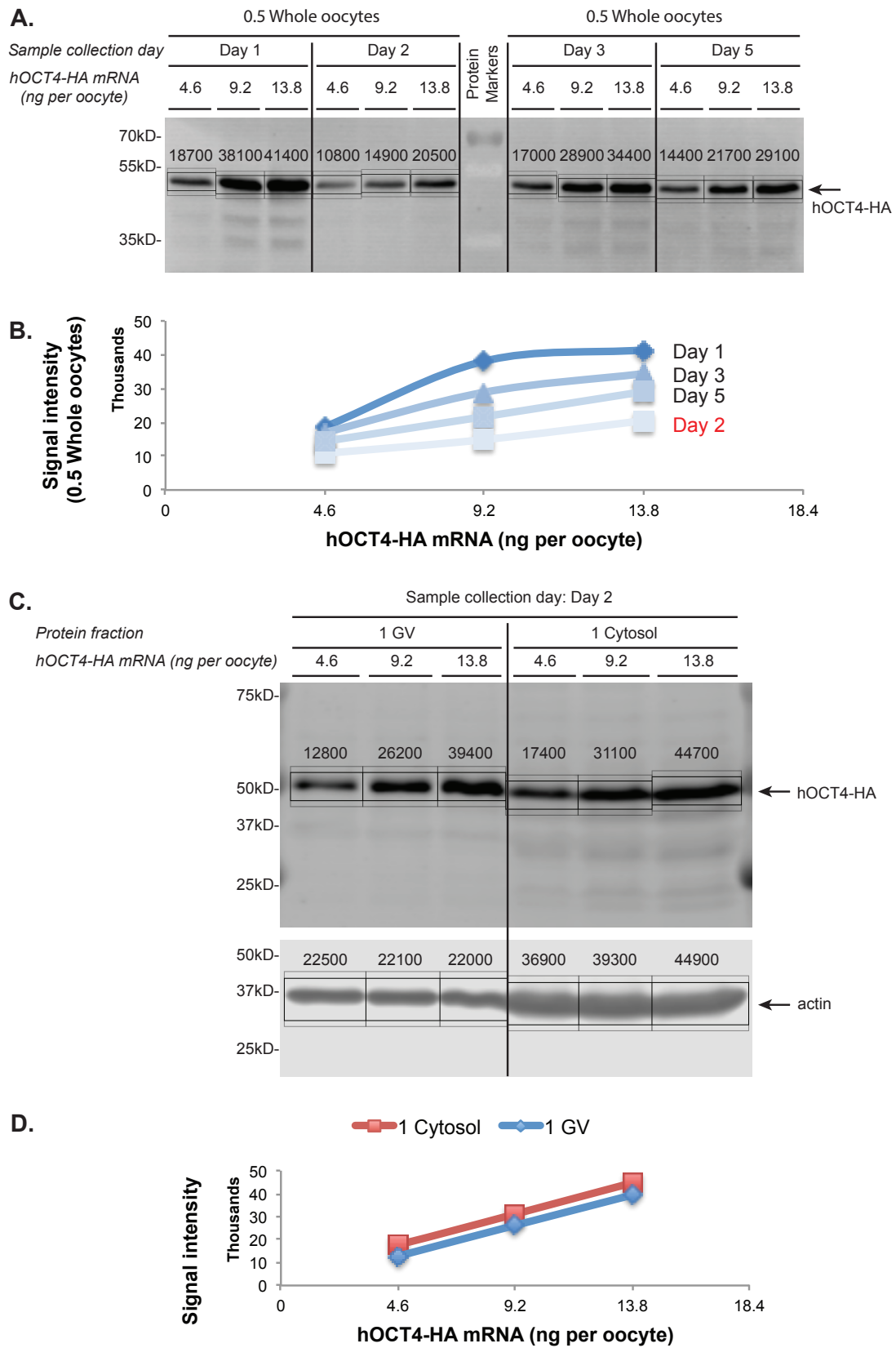
fractions and each lane is loaded with one GV fraction or one cytosol fraction for the Western blot. Black arrows indicate the positions of hOCT4-HA protein bands and actin protein bands. The signal intensity of hOCT4-HA protein bands is shown above the rectangles containing the quantification areas.

Comparing the signal intensity of hOCT4-HA proteins for 1 GV fraction to the signal intensity of hOCT4-HA proteins for 1 cytosol fraction, the signal intensity for 1 cytosol fraction is more than the signal intensity for 1 GV fraction in all three test doses (Figure 3.2.1.C-D). Taking the dose of 13.8 ng mRNA for example, the signal intensity for 1 cytosol fraction is 44700 and is more than 39400 for the signal intensity in 1 GV fraction (Figure 3.2.1.C). For the line graph demonstrating the correlation of the signal intensity of hOCT4-HA proteins versus hOCT4-HA mRNA (Figure 3.2.1.D), the nearly identical slopes of the lines between 1 cytosol fraction ($y=3x+4$) and 1 GV fraction ($y=2.9x-0.5$) show that the more hOCT4-HA mRNA injected into the cytosol of *Xenopus* oocytes, the more proteins made in the cytosol and the more proteins are transported into GV.

Considering the volume of cytosol is about 30 times more than the volume of GV in the *Xenopus* oocytes and the signal intensity of hOCT4-HA proteins in the cytosol fraction is almost equal to the signal intensity in the GV fraction (Figure 3.2.1.C), the hOCT4-HA proteins accumulate in the GV and the OCT4-HA proteins are actively transported into GV after being made in the cytosol from the injected hOCT4-HA mRNA.

Therefore, the amount of hOCT4-HA protein accumulating in the GVs increases when injecting the *Xenopus* oocytes with more hOCT4-HA mRNA. The more hOCT4-HA proteins accumulate in the GV, the greater is the chance hOCT4-HA proteins will bind to target sites to regulate downstream genes during the period after Oocyte-NT.

Figure 3.2.1



(Figure legend is on the next page)

Figure 3.2.1 hOCT4-HA proteins are produced dose-dependently and the amount of hOCT4-HA protein decreases at Day 2 or Day 3 after mRNA injection in the *Xenopus* oocytes. Additionally, hOCT4-HA proteins accumulate in the GV dose-dependently after hOCT4-HA proteins are made in the cytosol from the mRNA injected into the *Xenopus* oocytes

(A) The hOCT4-HA proteins of samples collected at different days after injecting the oocytes with different doses of mRNA are detected by anti-HA antibody and analysed on the Western blot. The position of hOCT4-HA protein bands is indicated (black arrow) and the signal intensity of hOCT4-HA protein bands is shown next to the rectangles containing the quantification areas.

(B) The line graph displays the dose-dependency of hOCT4-HA protein production after injecting the oocytes with different doses of hOCT4-HA mRNA at all sample collection days. The amount of hOCT4-HA protein decreases at Day 2 or Day 3 after mRNA injection. Since the signal intensity of each dose of Day 2 samples is less than the Day 3 and Day 5 samples with corresponding doses, the fall-off of hOCT4-HA signal of Day 2 samples is likely to be caused by incomplete protein transfer or improper sample preparation

(C) A Western blot of hOCT4-HA proteins from the samples collected at Day 2 after injecting the *Xenopus* oocytes with three doses of mRNA. Each lane is loaded with 1 cytosol fraction or 1 GV fraction. Black arrows indicate the position of hOCT4-HA proteins and actin and the signal intensity of hOCT4-HA proteins and actin are noted above the rectangles for quantification. For the signal intensity of 4.6 ng groups, the ratio of cytosol to GV is 1.4; for the signal intensity of 9.2 ng groups, the ratio of cytosol to GV is 1.2; for the signal intensity of 18.4 ng groups, the ratio of cytosol to GV is 1.1. Considering the ratio of the volume of cytosol to GV is about 30, hOCT4-HA proteins accumulate strongly in the GV after being made in the cytosol from injected mRNA.

(D) The line graphs are made by comparing the signal intensity of hOCT4-HA to the doses of hOCT4-HA mRNA. The equation of the line for the 1 cytosol groups is $y=3x+4$ and the slope is 3. The equation of the line for the 1 GV groups is $y=2.9x-0.4$ and the slope is 2.9. The nearly identical slope demonstrates that the more hOCT4-HA mRNAs are injected into the *Xenopus* oocytes, the more proteins are made in the cytosol and the more proteins are transported into GV.

3.2.1.2 Mouse Sox2-HA

3.2.1.2.1 The amount of mSox2-HA protein increases dose- and time-dependently after mRNA injection

To evaluate the production of mSox2-HA proteins, three doses of mSox2-HA mRNA are chosen, namely 4.6, 9.2 and 18.4 ng (Figure 3.2.2.A). The samples are collected at Day 1, 2 and 3 after mSox2-HA mRNA injection into *Xenopus* oocytes. Each lane of the Western blot is loaded with 0.7 whole oocytes. Black arrows indicate the positions of mSox2-HA protein bands and actin protein bands. The signal intensity of mSox2-HA proteins is shown next to the rectangles containing the quantification areas.

Comparing the samples collected at different days, the signal intensity of mSox2-HA per 0.7 whole oocytes increases over time from Day 1 to Day 3 after injecting each *Xenopus* oocyte with 9.2 ng of mSox2-HA mRNA (Figure 3.2.2.A-B). Comparing the samples at different doses of mSox2-HA mRNA, the signal intensity of mSox2-HA proteins per 0.7 oocytes increases from 4.6 ng to 18.4 ng of mSox2-HA mRNA in samples collected at Day 2 after mRNA injection (Figure 3.2.2.A and C). Since the increase of mSox2-HA proteins from 9.2 ng to 18.4 ng is less than the increase of mSox2-HA proteins from 4.6 ng to 9.2 ng, the production of mSox2-HA proteins is close to the plateau at 18.4 ng.

The production of mSox2-HA increases continuously from Day 1 to Day 3 after injecting the *Xenopus* oocytes with 9.2 ng of mSox2-HA mRNA and

therefore it would be fine to utilize this time period and mRNA dose for Oocyte-NT.

3.2.1.2.2 mSox2-HA proteins accumulate in the GV dose-dependently after mRNA injection

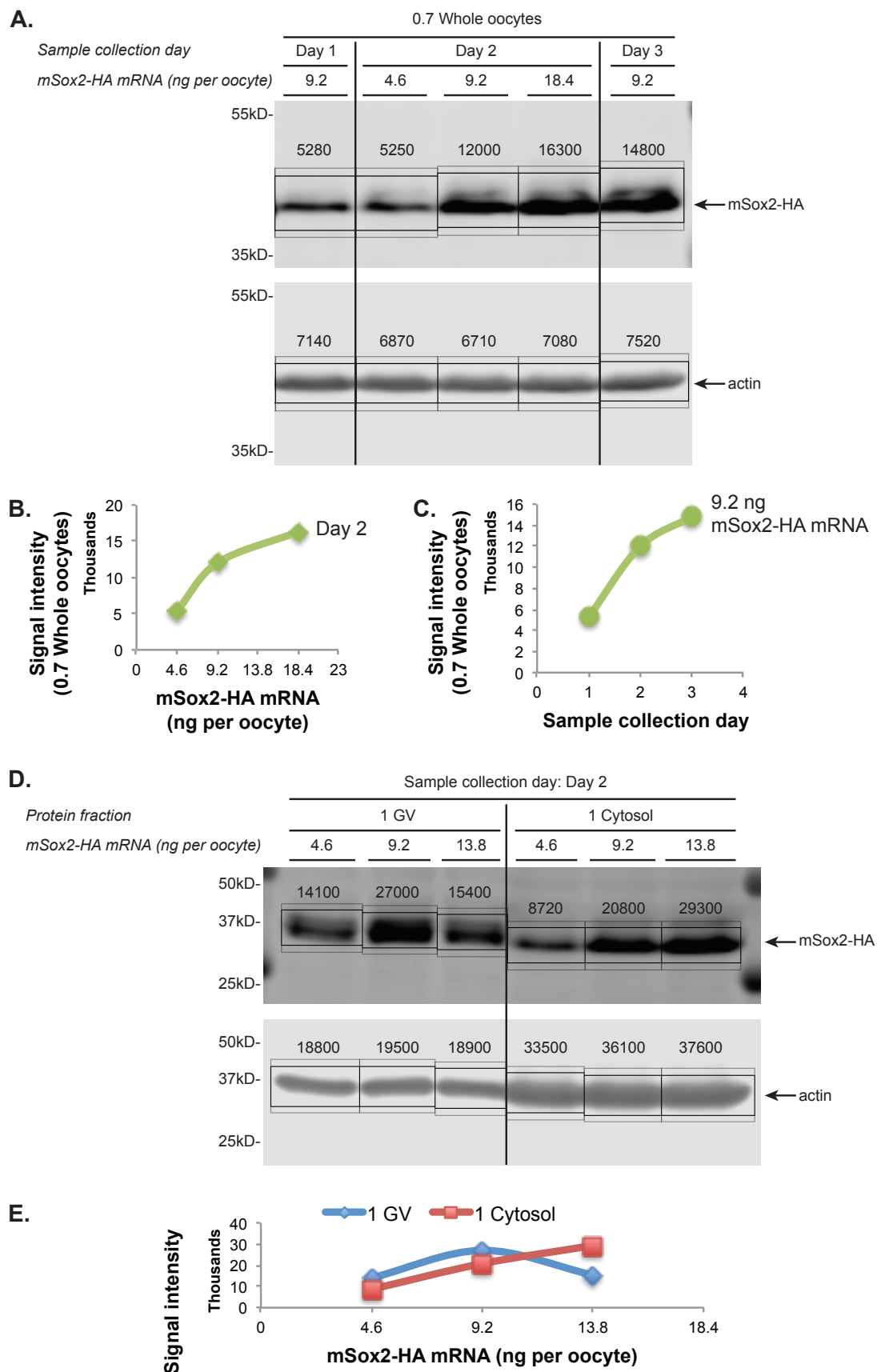
To evaluate the localization of mSox2-HA proteins after being produced in the cytosol from injected mSox2-HA mRNA, the samples are collected at Day 2 after mRNA injection and three doses of mRNA are used, namely 4.6 ng, 9.2 ng and 13.8 ng (Figure 3.2.2.D). The GV and cytosol fractions are separated by oocyte dissection and each lane of Western blot is loaded with 1 GV fraction or 1 cytosol fraction. The black arrows indicate the positions of mSox2-HA protein bands and actin protein bands. The signal intensity of mSox2-HA proteins are shown next to the black rectangles containing the areas for quantification.

Comparing the signal intensity of mSox2-HA proteins in GV samples at different doses (Figure 3.2.2.D-E), the signal intensity increases from 14100 for 4.6 ng mRNA group to 27000 for 9.2 ng mRNA group but decreases to 15400 for 13.8 ng mRNA group. The decrease of mSox2-HA protein from 9.2 ng mRNA group to 13.8 ng mRNA group may be caused by an inhibitory mechanism on the transportation of mSox2-HA protein to the GVs or caused by improper handling during oocyte dissection.

Although the amount of mSox2-HA protein falls at 13.8 ng (Figure 3.2.2.D-E), the signal intensity of mSox2-HA proteins at 4.6 ng and 9.2 ng still shows the accumulation of mSox2-HA proteins in the GVs after injecting the oocytes with mRNA when comparing the signal intensity of 1 cytosol fraction to the signal intensity of 1 GV fraction. Because the ratio of volume of cytosol to GV is 30:1 and the mSox2-HA protein signals of 1 GV fraction is more than the signal of 1 cytosol fraction for 4.6 ng groups and 9.2 ng groups, the mSox2-HA proteins obviously accumulate in the GV and the amount of mSox2-HA protein is more than the amount of mSox2-HA protein in the cytosol.

The dose-dependent accumulation of mSox2-HA proteins in the GV shows that the more mSox2-HA mRNAs are injected into oocytes, the more mSox2-HA proteins are synthesized in the cytosol and the more mSox2-HA proteins are transported into GV where the permeabilized MEFs will be transplanted into during Oocyte-NT. Interestingly, the difference in transportation between hOCT4-HA proteins (Figure 3.2.1) and mSox2-HA proteins (Figure 3.2.2) demonstrate the different mobility of different transcription factors and it relates to the binding of transcription factors to targets sites or the post-translational modification^{138,139}.

Figure 3.2.2



(Figure legend is on the next page)

Figure 3.2.2 The amount of mSox2-HA protein increases time- and dose-dependently and mSox2-HA proteins accumulate in the GV at Day 2 dose-dependently from 4.6 ng to 9.2 ng after mRNA injection.

(A) On the Western blot, the samples are collected at different days after injecting the *Xenopus* oocytes with different doses of mSox2-HA mRNA. The positions of mSox2-HA protein bands and actin protein bands are indicated (black arrows). The signal intensity of mSox2-HA proteins are shown next to the quantification areas (black rectangles).

(B) The signal intensity of mSox2-HA increases from Day 1 to Day 3 after injecting the oocytes with 9.2 ng of mSox2-HA mRNA.

(C) The signal intensity of mSox2-HA in the Day 2 samples increases from 4.6 ng to 18.4 ng of mSox2-HA mRNA. For the signal intensity of mSox2-HA, the ratio of 9.2 ng to 4.6 ng is 2.3; the ratio of 18.4 ng to 9.4 is 1.4. Since the doses double from 4.6 ng to 9.2 ng and from 9.2 ng to 18.4 ng, the increase of mSox2-HA proteins from 9.2 ng to 18.4 ng is close to the plateau due to the ratio of 18.4 ng to 9.4 ng is close to 1.

(D) On the Western blot, the signal intensity of mSox2-HA proteins and actin proteins are shown as numbers next to the quantification areas (black rectangles). The oocytes were injected with three different doses of mSox2-HA mRNA and samples were collected at Day 2 after mRNA injection. Each lane is loaded with the samples of 1 GV fraction or 1 cytosol fraction. The positions of mSox2-HA protein bands and actin protein bands are indicated (black arrows).

(E) The line graph shows the signal intensity of mSox2-HA proteins comparing the samples of 1 GV fraction and 1 cytosol fraction at three test doses of mRNA. The signal intensity of 1 GV fraction is more than 1 cytosol fraction at 4.6 ng and 9.2 ng but is less than 1 cytosol fraction at 13.8 ng. The fall of the mSox2-HA protein signals at 13.8 ng may be due to inhibitory mechanism on the transportation of mSox2-HA proteins or improper sample handling during oocyte dissection.

3.2.1.3 Human KLF4

3.2.1.3.1 hKLF4 proteins of GV fractions are detected by anti-KLF4 antibody and indicate the production and transportation of hKLF4 proteins after mRNA injection

Since there is no HA tag in the hKLF4 expression construct, an antibody (Abcam, ab194750) against human KLF2 was used for detecting the hKLF4 proteins for the 100% homology of the immunogenic sequence between hKLF2 and hKLF4. Additionally, the antibody is suggested to be able to recognize *Xenopus* klf2 (xklf2) and klf4 (xklf4) for the homology of the immunogenic sequence among different species. The exact immunogenic sequence is proprietary and it is a gift from Abcam (Data sheet is in appendices).

To evaluate the production and localization of hKLF4 proteins after injecting the oocytes with hKLF4 mRNA, Western blotting is performed (Figure 3.2.3.A). The samples are collected at Day 1 and Day 2 after injecting the oocytes with or without 9.2 ng of hKLF4 mRNA. The oocytes without mRNA injection are also collected because the antibody is suggested to recognize xklf2 and xklf4 and the endogenous xklf2 and xklf4 can be distinguished by comparing the no mRNA injection groups and hKLF4 mRNA groups. Each lane of the Western blot is loaded with 1.5 whole oocytes, 1.5 cytosol fractions, 18 GV fractions or 36 GV fractions. The predicted positions of hKLF4 proteins and the position of actin proteins are indicated by black arrows. The signal intensity of hKLF4

protein bands and actin protein bands is shown next to the black rectangles containing the quantification areas.

Comparing the GV fractions of no mRNA injection group to the GV fractions of hKLF4 mRNA group (Figure 3.2.3.A), hKLF4 proteins are recognized by antibody in the mRNA injection group because the signal intensity of band d for 18 GV fractions (signal intensity=2510) is 14 times more than the signal intensity of band a for 36 GV fractions (signal intensity=179) at the predicted hKLF4 position 1. Since the antibody can recognize *xklf2* and *xklf4*, the four visible bands, b, c, e and f, are most likely to be endogenous *xklf2/4* protein or unspecific bands (Figure 3.2.5.A and B). The reason is that the signal intensity for 18 GV fractions of hKLF4 mRNA group (Band e at Position 2 and Band f at Position 3) is less than the signal intensity for 36 GV fractions of no mRNA injection group (Band b at Position 2 and Band c at Position 3) while the number of GV fractions in the no mRNA injection group is twice as many as the number of GV fractions in the hKLF4 mRNA group (Figure 3.2.3).

When I tried to locate the predicted hKLF4 positions, I have found that the relatively more mineral oil in the GV fractions than the mineral oil in the cytosol fractions makes the actin protein bands of GV fractions shift upwards on the blots, compared to the actin protein bands of the cytosol fractions (white arrowhead, Fig 3.2.3.A). Therefore, the positions of bands representing the same proteins may differ if the samples contain different levels of mineral oil and it needs extra care when performing oocyte dissection and analyzing the samples through Western blotting.

Although hKLF4 proteins can be detected in the GV fractions, no visible protein bands of the cytosol fractions and of whole oocytes are seen on the blot at the predicted hKLF4 positions or proximity at Day 1 and Day 2 after mRNA injection (Figure 3.2.3.A and 3.2.4.A). When quantifying the hKLF4 protein signals of whole oocytes and the cytosol fractions at predicted hKLF4 positions (Figure 3.2.4), the signal intensity is below zero since the signals in the quantification areas (middle rectangles) reduce the signals in the background areas (top and bottom black rectangles surround the middle black rectangles) for normalization. Comparing the hKLF4 signals of 1 whole oocyte, 1 cytosol fraction and 1 GV fraction (Figure 3.2.4.B), the hKLF4 signals are 131 for 1 GV fraction of hKLF4 mRNA group and 5 for 1 GV fraction of no mRNA injection group whereas the signals for 1 whole oocyte and 1 cytosol fraction are below zero of both groups (Figure 3.2.4.B).

Although the specificity of hKLF4 antibody is not enough to recognize the hKLF4 proteins in whole oocytes and the cytosol fractions, the signals of hKLF4 proteins in the GV fractions indicate that hKLF4 proteins can be synthesized from hKLF4 mRNA and transported into GV. This is an important check before performing Oocyte-NT. In addition to this point, it will demonstrate that the overexpression of hKLF4 proteins has effects on the transplanted MEFs and regulates certain pluripotency genes in 3.2.2.3.

Figure 3.2.3

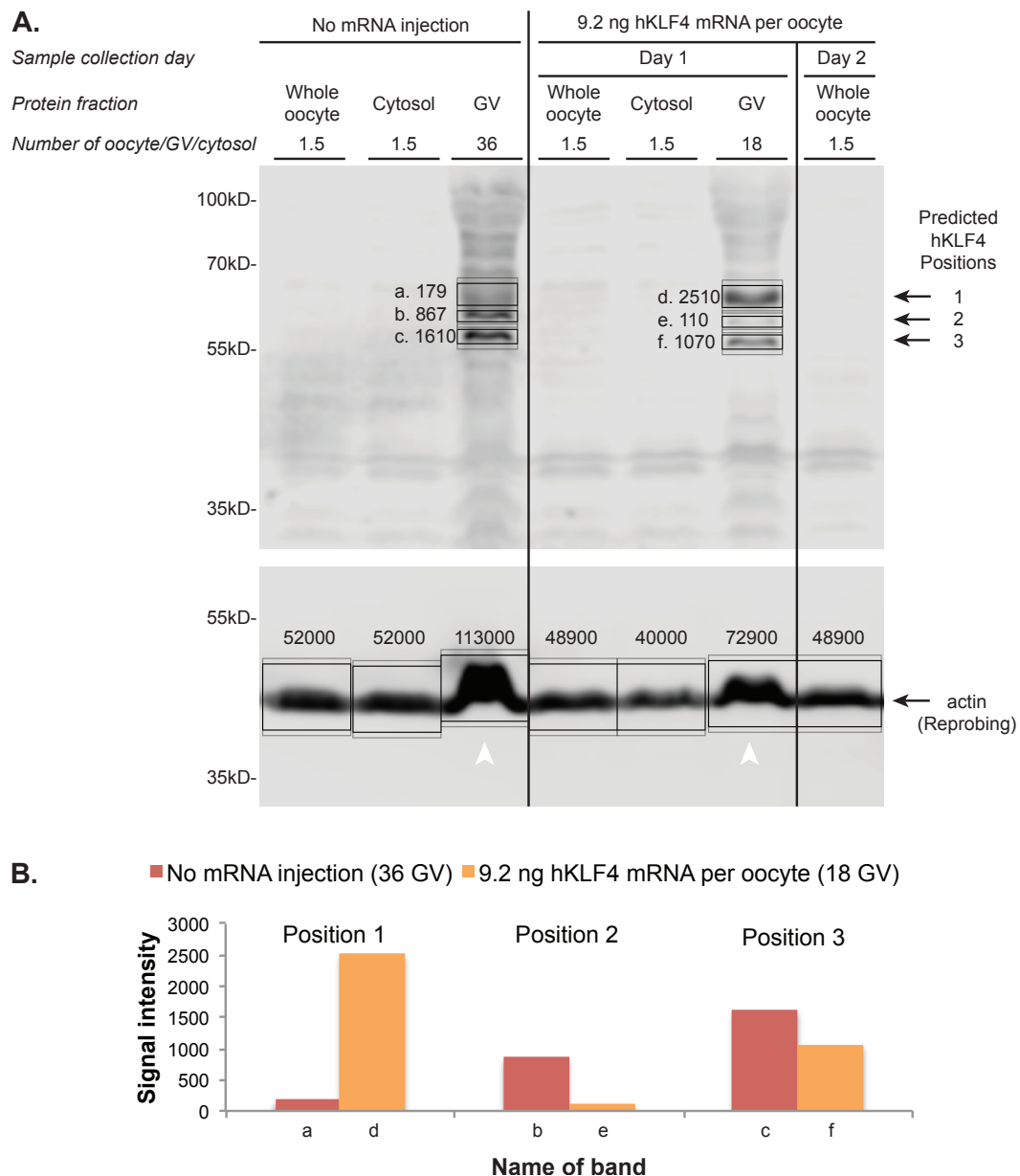


Figure 3.2.3 hKLF4 protein band at position 1 and unknown protein bands at position 2 and 3 are detected in the GV fractions, but not detected in whole oocytes and the cytosol fractions.

(A) The samples of hKLF4 mRNA groups are collected at Day 1 and Day 2 after mRNA injection. In the GV fractions, three positions are predicted for hKLF4 protein bands (black arrows, ~58kD) and visible bands are detected by the anti-hKLF2/4 antibody in no mRNA injection groups (band a, b and c) and hKLF4 mRNA groups (band d, e and f). Actin protein bands (black arrow) are reprobated after hKLF4 detection because antibodies against actin and hKLF4 are raised from rabbit and hKLF4, xklf2/4 and actin have similar sizes.

(Figure legend continues on the next page)

The actin bands of GV fractions shift upwards (white arrows) since the mineral oil in the GV fractions is relatively higher than in the cytosol fractions. Signal intensity of hKLF4 proteins and actin proteins is shown next to the quantification area (black rectangles).

(B) Because 36 GV fractions of no mRNA injection group is more than 18 GV fractions of hKLF4 mRNA group, the endogenous xklf2/4 protein signals or unspecific protein signals of hKLF4 mRNA groups are expected to be lower than the signals of no mRNA injection groups. On the contrary, if the protein signals of the hKLF4 mRNA group are higher than the signals of no mRNA injection group, the bands represent where hKLF4 proteins locate. Therefore, Position 1 indicates the position of hKLF4 protein bands and Position 2 and 3 indicate endogenous xklf2/4 or unspecific protein bands.

Figure 3.2.4

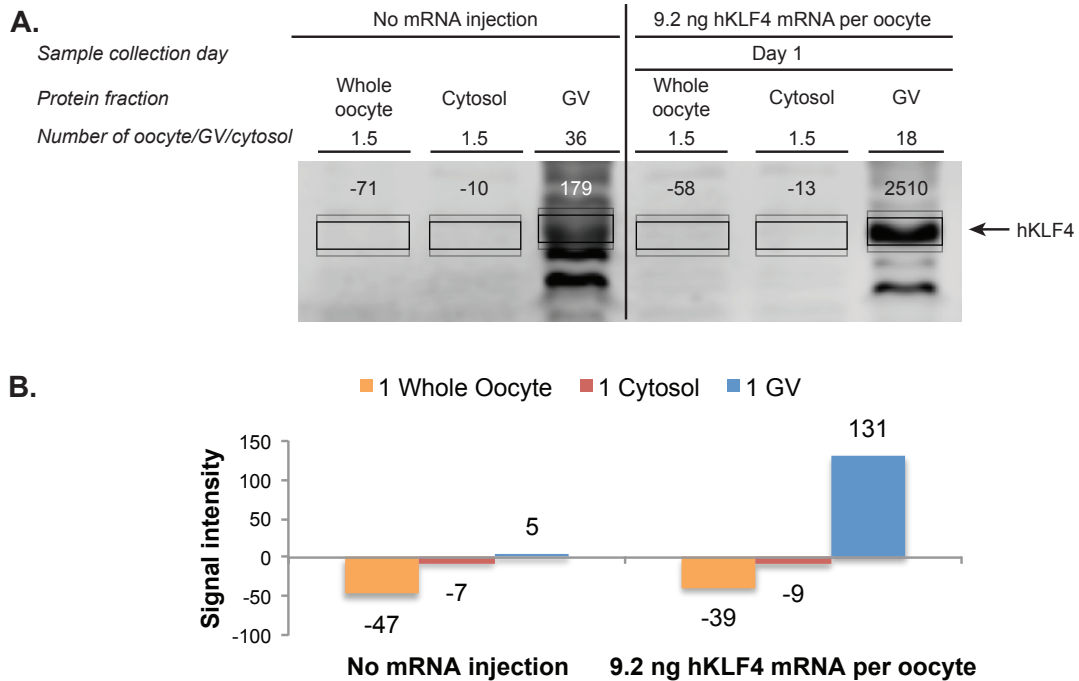


Figure 3.2.4 No hKLF4 signals are detected in whole oocytes and the cytosol fractions at the position of hKLF4 proteins.

(A) The magnified image of Fig 3.2.5.A. No visible bands are seen in the samples of whole oocyte or the cytosol fractions of hKLF4 mRNA injection groups. The black arrow indicates Position 1 of predicted hKLF4 positions. The signal intensity quantified at hKLF4 Position 1 is shown next to the quantification areas (black rectangles) for the samples of whole oocytes, the cytosol fractions and GV fractions.

(B) The signals at hKLF4 Position 1 for 1 whole oocyte, 1 cytosol fraction and 1 GV fraction are compared. It shows no signals of 1 whole oocyte or 1 cytosol fraction are detected in no mRNA injection and hKLF4 mRNA groups since they are all below zero.

3.2.2 The overexpression of mammalian YFs regulates pluripotency genes in MEFs at Day 2 after Oocyte-NT

In 3.2.1, I have confirmed that the proteins of mammalian Yamanaka factors (YFs) can be made in the *Xenopus* oocytes from the corresponding mRNA and the proteins made in the cytosol of oocytes can be transported into the GV. In 3.2.2, I have performed Oocyte-NT after overexpressing each YF for one day and evaluated how the overexpression of each YF affects the expression of selected pluripotency genes of MEFs under the process of somatic cell nuclear reprogramming (SCNR) by *Xenopus* oocytes.

The experimental settings are the same for each YF. Twenty-four hours before Oocyte-NT, 9.2 ng mRNAs of each YF are injected into the cytosol of *Xenopus* oocytes. The immortalized mouse embryonic fibroblast (MEFs) strain, sixiFM, is used in Oocyte-NT. At Day 0, Oocyte-NT is performed and SLO-permeabilized MEFs are transplanted into the GVs of oocytes (~500 MEFs per oocyte). The oocyte samples are collected two days after Oocyte-NT. The RNA of oocytes will be processed and analyzed by QPCR.

Two groups are compared for each YF factor. In the control groups, SLO-permeabilized MEFs are transplanted into GVs at Day 0 without mRNA injection. In YF overexpression groups, the mRNA injection of each YF is performed 24 hours before Oocyte-NT and then SLO-permeabilized MEFs are transplanted into GVs at Day 0. To quantify the expression of pluripotency genes of MEFs by QPCR, the relative amount of transcript of mGAPDH is

used as an internal control for normalizing the relative amount of transcript of pluripotency genes and also for estimating the number of transplanted MEFs.

The effect of YF on pluripotency genes is defined by the fold change between the YF overexpression groups and control groups. If the average fold change is more than 2 or less than 0.5, pluripotency gene is judged to be up-regulated or down-regulated by the overexpression of one of the YFs. The biological replicates are produced from different batches of *Xenopus* oocytes and each batch of oocytes is taken from the ovary of different female frog. The variable fold change is caused by the individual difference between *Xenopus* oocytes and the resistance of MEF genes to the maternal factors and the overexpressed transcription factors and this point will be discussed further in the following chapters.

3.2.2.1 The overexpression of hOCT4-HA up-regulates the expression of mSox2 in MEFs at Day 2 after Oocyte-NT

To evaluate the effect of hOCT4-HA proteins on pluripotency genes during SCNR by oocytes, Oocyte-NT is performed with SLO-permeabilized MEFs after overexpressing the hOCT4-HA proteins for one day. The result shows that the overexpression of hOCT4-HA up-regulates the expression of mSox2 by 3.6-fold and does not regulate 8 other pluripotency genes with average fold changes less than 2 and more than 0.5 (Figure 3.2.5.A).

3.2.2.2 The overexpression of mSox2-HA up-regulates the expression of mOct4, mSox2 and mKlf4 in MEFs at Day 2 after Oocyte-NT

To evaluate the effect of mSox2-HA proteins on pluripotency genes during SCNR by oocytes, Oocyte-NT is performed with SLO-permeabilized MEFs after overexpressing the mSox2-HA proteins for one day. The result shows that the overexpression of mSox2-HA up-regulates the expression of mOct4 by 2.6-fold, mSox2 by 14.4-fold and mKlf4 by 2.3-fold (Figure 3.2.5.B).

3.2.2.3 The overexpression of hKLF4 up-regulates the expression of mOct4, mSox2, mKlf2, mSall4 and mUtf1 in MEFs at Day 2 after Oocyte-NT

To evaluate the effect of hKLF4 proteins on pluripotency genes during SCNR by oocytes, Oocyte-NT is performed with SLO-permeabilized MEFs after overexpressing the hKLF4 proteins for one day. The result shows that the overexpression of hKLF4 up-regulates the expression of mOct4 by 3.3-fold ($p<0.05$), mSox2 by 4.3-fold, mKlf2 by 1.3-fold ($p<0.01$), mSall4 by 4-fold ($p<0.05$) and mUtf1 by 2.9-fold ($p<0.05$) in Figure 3.2.5.C.

Figure 3.2.5

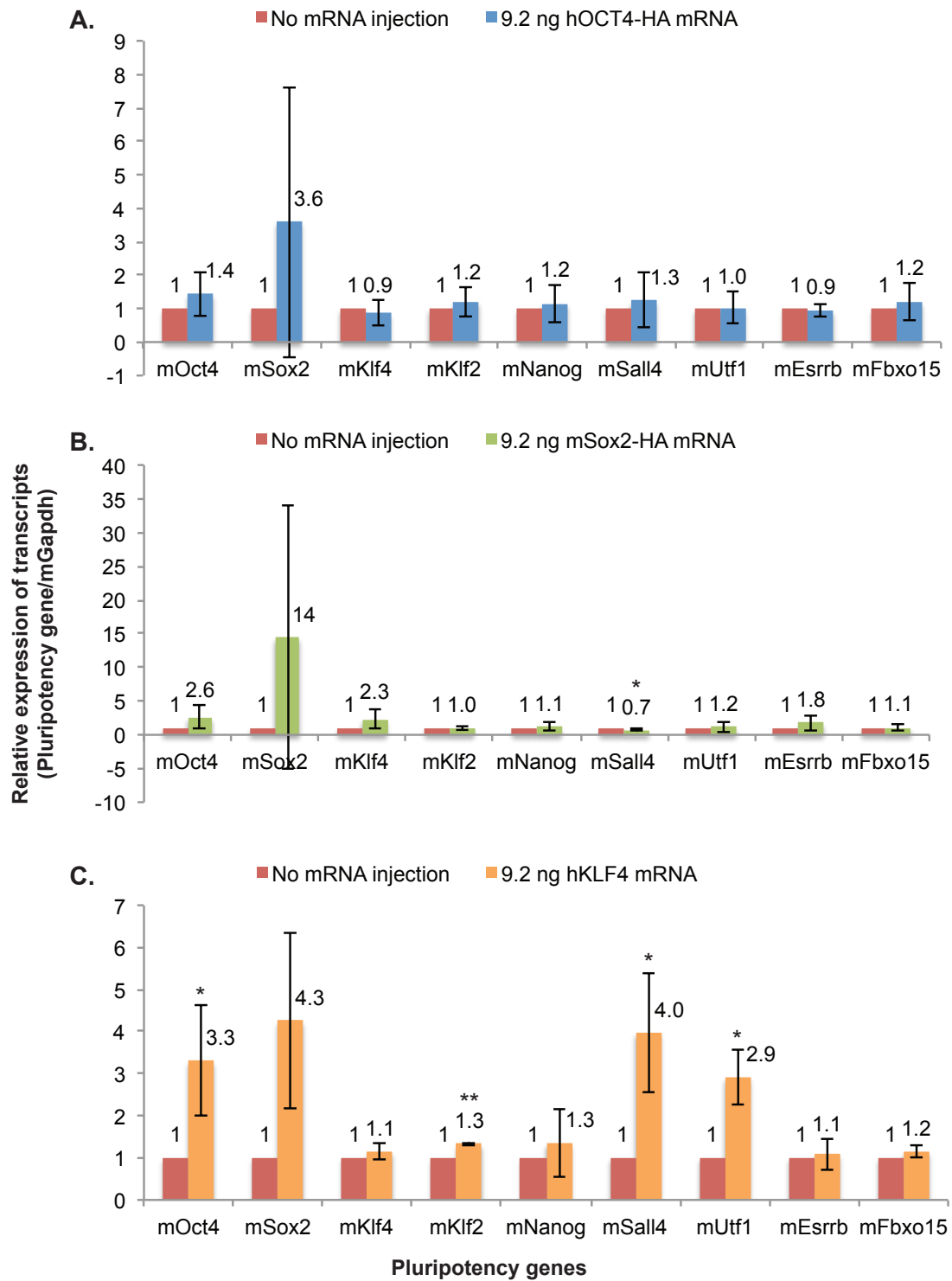


Figure 3.2.5 The overexpression of mammalian Yamanaka homologs up-regulates different sets of selected pluripotency genes in MEFs at Day 2 after Oocyte-NT

The relative expression of transcripts in control group (No mRNA injection) is 1 for each pluripotency gene and shown as a bar in red. (Figure legend continues on the next page)

(A) The relative expression of transcripts in hOCT4-HA overexpression group (9.2 ng hOCT4-HA mRNA) for each pluripotency gene represents the effect of the overexpression of hOCT4-HA proteins and shown as bars in blue (n=3, *t*-test).

(B) The relative expression of transcripts in mSox2-HA overexpression group (9.2 ng mSox2-HA mRNA) for each pluripotency gene represents the effect of the overexpression of mSox2-HA proteins and shown as bars in green (**p*<0.05, n=3, *t*-test).

(C) The relative expression of transcripts in hKLF4 overexpression group (9.2 ng hKLF4 mRNA) for each pluripotency gene represents the effect of the overexpression of hKLF4 proteins and shown as bars in orange (**p*<0.05, ***p*<0.01, n=2, *t*-test).

3.2.3 Summary

To sum, proteins of mammalian YF homologs, hOCT4-HA and mSox2-HA, can be synthesized in cytosol of *Xenopus* oocytes from injected mRNA dose-dependently and time-dependently. After YF proteins are made, these proteins are transported into GV and accumulate in the GV dependent on the dose of mRNA injected. Therefore, dose- and time- dependent synthesis of mammalian proteins with HA tag can be measured in *Xenopus* oocytes via our Western blotting system.

Additionally, the regulation of pluripotency genes in transplanted nuclei by mammalian YF homologs can be detected in the *Xenopus* oocyte system after nuclear transfer although the protein degradation of hOCT4-HA is observed at Day 2 after mRNA injection and hKLF4 can not be detected due to the lack of HA tag and useful antibody.

It shows that different mammalian YFs regulate different sets of tested pluripotency genes in sixiFM MEFs. hOCT4-HA up-regulates the expression of mSox2 by 3.6-fold at Day 2 after Oocyte-NT. mSox2-HA up-regulates the expression of mOct4 by 2.6-fold, mSox2 by 14-fold and mKlf4 by 2.3-fold. hKLF4 up-regulates the expression of mOct4 by 3.3-fold ($p<0.05$), mSox2 by 4.3-fold, mSall4 by 4-fold ($p<0.05$) and mUtf1 by 2.9-fold ($p<0.05$).

3.3 *Xenopus* Yamanaka factors can be produced in the *Xenopus* oocytes after mRNA injection and regulate pluripotency genes in MEFs at Day 2 after Oocyte-NT

In 3.2, it has been proven that the proteins of mammalian Yamanaka factors (YFs) can be produced in the cytosol of the *Xenopus* oocytes from injected mRNA and transported into GV where the YFs bind to their target genes. It has also been shown that the overexpression of each mammalian YF can regulate different sets of pluripotency genes in MEFs during the somatic cell nuclear reprogramming (SCNR) by oocytes.

The next question is whether *Xenopus* YFs can regulate pluripotency genes in MEFs through a cross-species interaction? Therefore, I chose three *Xenopus* homologs of YFs, namely *xklf2*-HA, *xpou60* and *xsox2*, and examine their ability to regulate pluripotency genes in MEFs.

To evaluate the effect of *Xenopus* YFs on pluripotency genes in MEFs during SCNR by oocytes, Oocyte-NT is performed after overexpressing each *Xenopus* YF for one day. Because the binding of transcription factors to their target genes relates to the amount of transcription factor, the production and localization of *Xenopus* YFs after mRNA injection were examined via Western blotting before performing Oocyte-NT (Section 3.3.1). After confirming the *Xenopus* YF proteins can be produced in the cytosol and transported into GVs, Oocyte-NT is performed (Section 3.3.2).

3.3.1 The production and localization of *Xenopus* YFs after mRNA injection in *Xenopus* oocytes

To evaluate the production and localization of *Xenopus* YFs in the *Xenopus* oocytes, Western blotting was performed. Different doses of YF mRNA were injected into the cytosol of *Xenopus* oocytes at Day 0 and samples were collected at different time points after mRNA injection. The localization of YF proteins made in the cytosol from injected mRNA is evaluated by comparing the amount of YF protein in the GV fractions to the amount in whole oocytes. The YF proteins on the Western blots are detected by the anti-HA antibody and actin detected by anti-actin antibody is used as internal control. The amount of YF protein and actin proteins is quantified by LI-COR Odyssey® CLx Imaging System.

3.3.1.1 *Xenopus* sox2-HA

3.3.1.1.1 The amount of xsox2-HA protein increases dose- and time-dependently after mRNA injection

Comparing the signal intensity of xsox2-HA proteins at different sample collection days after injecting the oocytes with 9.2 ng xsox2-HA mRNA (Figure 3.3.1.A-B), the xsox2-HA proteins increase from Day 1 to Day 3 because the signal intensity of xsox2-HA per oocyte goes up from 32429 for Day 1 sample to 38143 for Day 3 sample. Comparing the signal intensity of Day 2 samples (Figure 3.3.1.A and C), the amount of xsox2-HA protein increase when injecting the oocytes with more xsox2-HA mRNA and the signal intensity of

xsox2-HA per oocyte increases from 22571 for 4.6 ng mRNA to 49857 for 18.4 ng mRNA.

In conclusion, the production of xsox2-HA proteins is dose-dependent from the dose of 4.6 ng to 18.4 ng xsox2-HA mRNA. In addition, the xsox2-HA proteins increase continuously from Day 1 to Day 3 after injecting the oocytes with 9.2 ng mRNA and this coincides with the time points when Oocyte-NT is performed and when the samples are collected.

3.3.1.1.2 xsox2-HA proteins accumulate in the GVs after mRNA injection

Comparing the signal intensity of 1 whole oocyte to the signal intensity of 1 GV fraction (Figure 3.3.1.A and D), the signal intensity is 35286 for 1 whole oocyte and 22938 for 1 GV fraction. Since the ratio of the volume of 1 whole oocyte to 1 GV fraction is ~30 and the ratio of xsox2-HA signals of 1 whole oocyte to 1 GV fraction is ~1.5, the xsox2-HA proteins accumulate in the GVs after the xsox2-HA proteins are made in the cytosols and transported into GVs. It means that 40% of xsox2-HA proteins produced in the cytosol will enter the GVs and therefore the concentration of xsox2-HA proteins in the GVs is about 20 times more than the concentration in whole oocytes.

The accumulation of xsox2-HA proteins in the GVs indicate the xsox2-HA proteins are transported into GVs where xsox2-HA proteins can bind to target genes as well as the genes in the MEFs via Oocyte-NT.

Figure 3.3.1

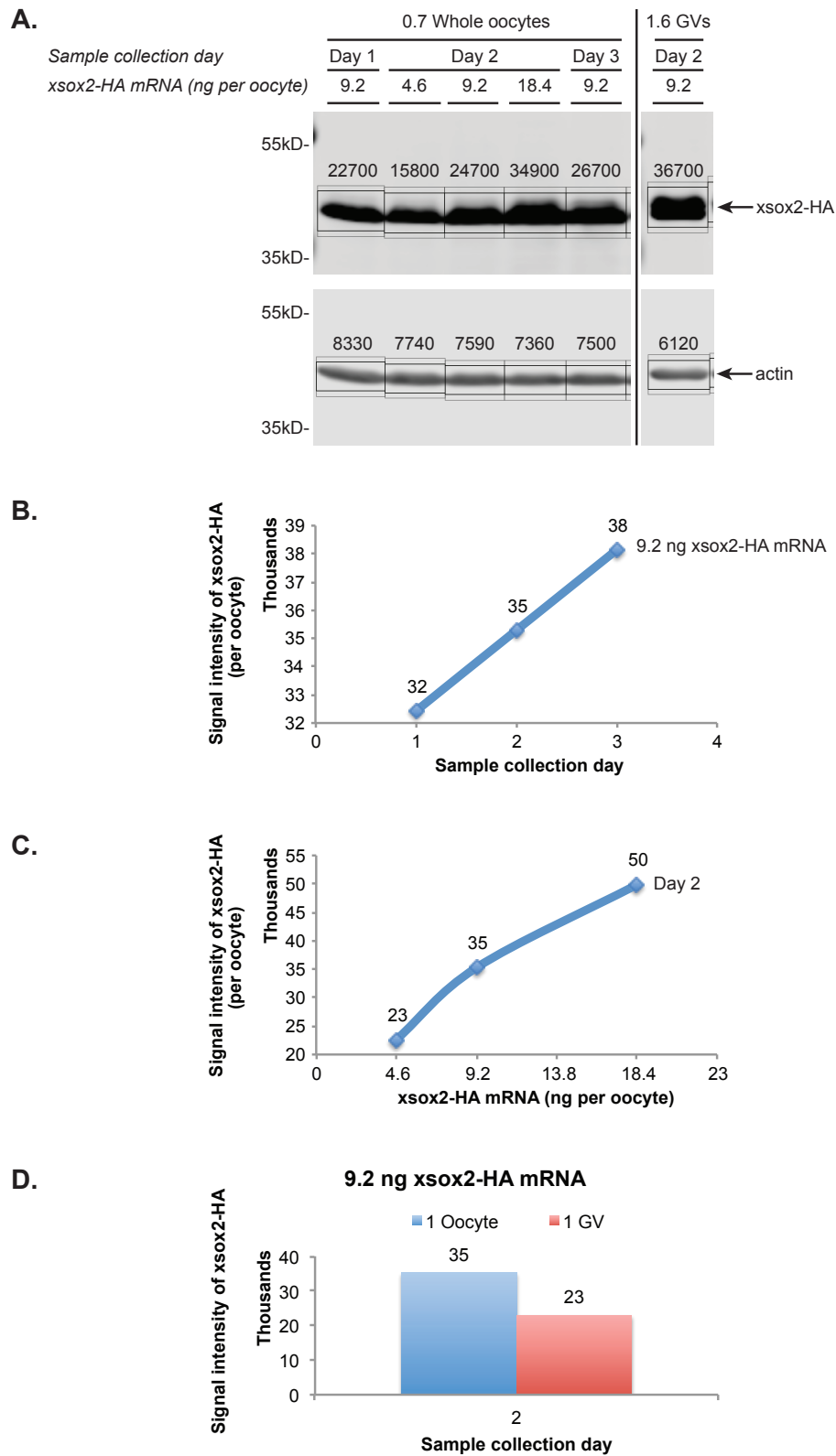


Figure 3.3.1 xsox2-HA proteins increase time- and dose-dependently and accumulate in the GV after mRNA injection. (Figure legend continues on the next page)

(A) The xsox2-HA proteins and actin proteins of different samples are detected on the Western blot. The positions of xsox2-HA and actin protein bands are indicated (black arrows). Signal intensity of xsox2-HA is noted next to the quantification areas (black rectangles).

(B) The production of xsox2-HA proteins is time-dependent. For the 9.2 ng xsox2-HA mRNA groups, the signal intensity of Day 1 sample is 32429 per oocyte; the signal intensity of Day 2 sample is 35286 per oocyte; the signal intensity of Day 3 sample is 38143 per oocyte.

(C) The production of xsox2-HA proteins is dose-dependent. For the Day 2 samples, the signal intensity of 4.6 ng group is 22571; the signal intensity of 9.2 ng group is 35286; the signal intensity of 18.4 ng group is 49857.

(D) The xsox2-HA proteins accumulate in the GVs after the xsox2-HA proteins are made in the cytosol and transported into GVs. For the 9.2 ng xsox2-HA mRNA groups of Day 2 samples, the ratio of signal intensity of 1 whole oocyte to 1 GV fraction is ~ 1.5 and it is $1/20$ of the ratio of volume of 1 whole oocyte to 1 GV fraction, which is ~ 30 . It means the concentration of xsox2-HA proteins in the GVs is 20 times as much as the concentration in the whole oocytes.

3.3.1.2 *Xenopus klf2-HA*

3.3.1.2.1 The amount of xklf2-HA increase dose- and time-dependently from Day 1 to Day 2 but decrease at Day 3 after mRNA injection

Comparing the signal intensity of xklf2-HA proteins at different days after injecting the oocytes with 9.2 ng xklf2-HA mRNA per oocyte (Figure 3.3.2.A-B), the signal intensity per oocyte increases as shown in Figure 3.3.2.B.

The decrease of xklf2-HA proteins may affect the effectiveness of xklf2-HA on regulating downstream genes if the inhibitory mechanism of *Xenopus* oocytes is involved but it is not observed in 3.3.2.3 when the samples are collected at Day 3 after mRNA injection.

3.3.1.2.2 The xklf2-HA proteins accumulate in the GVs dose-dependently after mRNA injection

Comparing the signal intensity of xklf2-HA proteins of Day 2 samples between 1 whole oocyte and 1GV fraction (Figure 3.3.2.A and C), it shows the xklf2-HA proteins accumulate in the GV because the ratio of the volume of 1 whole oocyte to 1 GV fraction is ~30 and the ratio of the signal intensity of 1 whole oocyte to 1 GV fraction is ~5 for both 9.2 ng and 18.4 ng mRNA groups. It means 1/6 of xklf2-HA proteins produced in the cytosol will enter the GVs and therefore the concentration of xklf2-HA proteins in the GVs is about 6 times as much as the concentration in whole oocytes.

Since the ratio of the xklf2-HA signals of 1 whole oocyte to 1 GV fraction is ~5 for both 9.2 ng and 18.4 ng mRNA groups, the more xklf2-HA proteins produced from the injected mRNA in the cytosol, the more xklf2-HA proteins will be transported into GV.

In sum, the accumulation of xklf2-HA proteins in the GV is dose-dependent. Therefore, the more xklf2-HA mRNAs are injected into *Xenopus* oocytes, the more xklf2-HA proteins will be made in the cytosol and the more xklf2-HA proteins will be in the GV to bind to their target genes during period after Oocyte-NT.

Figure 3.3.2

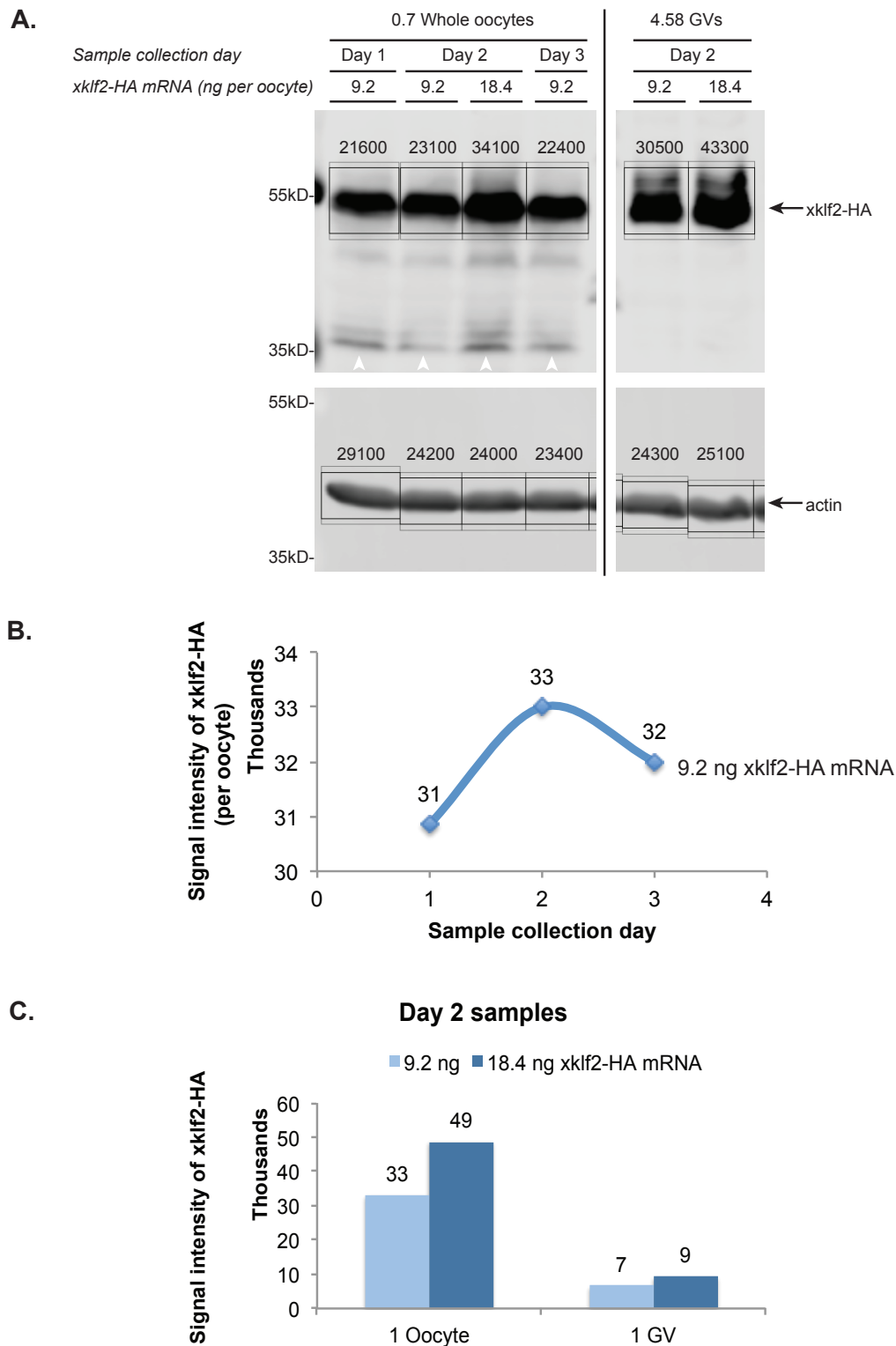


Figure 3.3.2 *xklf2*-HA proteins increase dose- and time-dependently from Day 1 to Day 2 but decrease at Day 3 after mRNA injection. *xklf2*-HA proteins accumulate in GVs after *xklf2*-HA proteins are made in cytosols from *xklf2*-HA mRNA.

(Figure legend continues on the next page)

(A) The xklf2-HA and actin proteins are detected on the Western blot and indicated by black arrows. Signal intensity of xklf2-HA and actin proteins are shown above the quantification areas (black rectangles). Some degradation of xklf2-HA proteins are observed in the whole oocytes but not in the GV fractions (white arrowheads).

(B) The time-dependent production of xklf2-HA proteins. For the 9.2 ng xklf2-HA mRNA groups, the signal intensity of Day 1 sample is 30857 per oocyte; the signal intensity of Day 2 sample is 33000 per oocyte; the signal intensity of Day 3 sample is 32000 per oocyte.

(C) The dose-dependent production of xklf2-HA proteins in 1 whole oocyte and the dose-dependent accumulation of xklf2-HA proteins in 1 GV fraction. For the Day 2 samples, the signal intensity of 1 whole oocyte is 33000 for 9.2 ng mRNA group and 48714 for 18.4 ng mRNA group; the signal intensity of 1 GV fraction is 6659 for 9.2 ng mRNA group and 9454 for 18.4 ng mRNA group. The ratio of signal intensity of 1 whole oocyte to 1 GV fraction is ~5 for 4.6 ng and 9.2 ng mRNA groups and it is 1/6 of ratio of volume of 1 whole oocyte to 1 GV fraction, which is ~30. It means the concentration of xklf2-HA proteins in the GVs is 6 times as much as the concentration in the whole oocytes.

3.3.2 The overexpression of *Xenopus* YFs regulates pluripotency genes in MEFs at Day 2 after Oocyte-NT

In subsection 3.3.1, it shows that the *Xenopus* YF proteins can be made from injected mRNA in the cytosol of *Xenopus* oocytes and then be transported into and accumulate in the GV. In subsection 3.3.2, I have performed Oocyte-NT after overexpressing each YF for one day and evaluate how the overexpression of each YF affects the expression of selected pluripotency genes in MEFs during the somatic cell nuclear reprogramming by *Xenopus* oocytes.

The experimental settings are the same as Oocyte-NT performed for mammalian YFs and described in subsection 3.2.2. Briefly, 9.2 ng of each *Xenopus* YF mRNA is injected into *Xenopus* oocytes a day before Oocyte-NT. The MEF cell line, sixiFM, is used for Oocyte-NT. The oocyte samples are collected two days after Oocyte-NT. The RNA of oocytes will be processed and analyzed by QPCR.

The normalization of transcripts and the definition of fold change have been described in subsection 3.2.2. In short, the relative amount of pluripotency gene is normalized by the relative amount of internal control, mGAPDH. The fold change between control group (MEF) and YF overexpression group (MEF+YF mRNA) defines if pluripotency genes are regulated by the overexpression of one of the *Xenopus* YFs. If the fold change of pluripotency gene is more than two or less than 0.5, this gene is up-regulated or down-regulated by the overexpression of YFs.

3.3.2.1 The overexpression of xpou60-HA upregulates the expression of mSox2, mKlf2 and mSall4 in MEFs at Day 2 after Oocyte-NT

To evaluate the effect of xpou60-HA proteins on pluripotency genes during SCNR by oocytes, Oocyte-NT is performed with SLO-permeabilized MEFs after overexpressing the xpou60-HA proteins for one day. The result shows that the overexpression of xpou60-HA up-regulates the expression of mSox2, mKlf2, and mSall4 as shown in Figure 3.3.3.A.

3.3.2.2 The overexpression of xsox2-HA does not regulate the selected pluripotency genes in MEFs at Day 2 after Oocyte-NT

To evaluate the effect of xsox2-HA proteins on pluripotency genes during SCNR by oocytes, Oocyte-NT is performed with SLO-permeabilized MEFs after overexpressing the xsox2-HA proteins for one day. The result shows that the overexpression of xsox2-HA does not regulate any of the selected pluripotency genes (Figure 3.3.3.B).

3.3.2.3 The overexpression of xklf2-HA upregulates the expression of mOct4, mSox2, and mSall4 in MEFs at Day 2 after Oocyte-NT

To evaluate the effect of xklf2-HA proteins on pluripotency genes during SCNR by oocytes, Oocyte-NT is performed with SLO-permeabilized MEFs after overexpressing the xklf2-HA proteins for one day. The result shows that the overexpression of xklf2-HA up-regulates the expression of mOct4, mSox2 and mSall4 as shown in Figure 3.3.3.C.

Figure 3.3.3

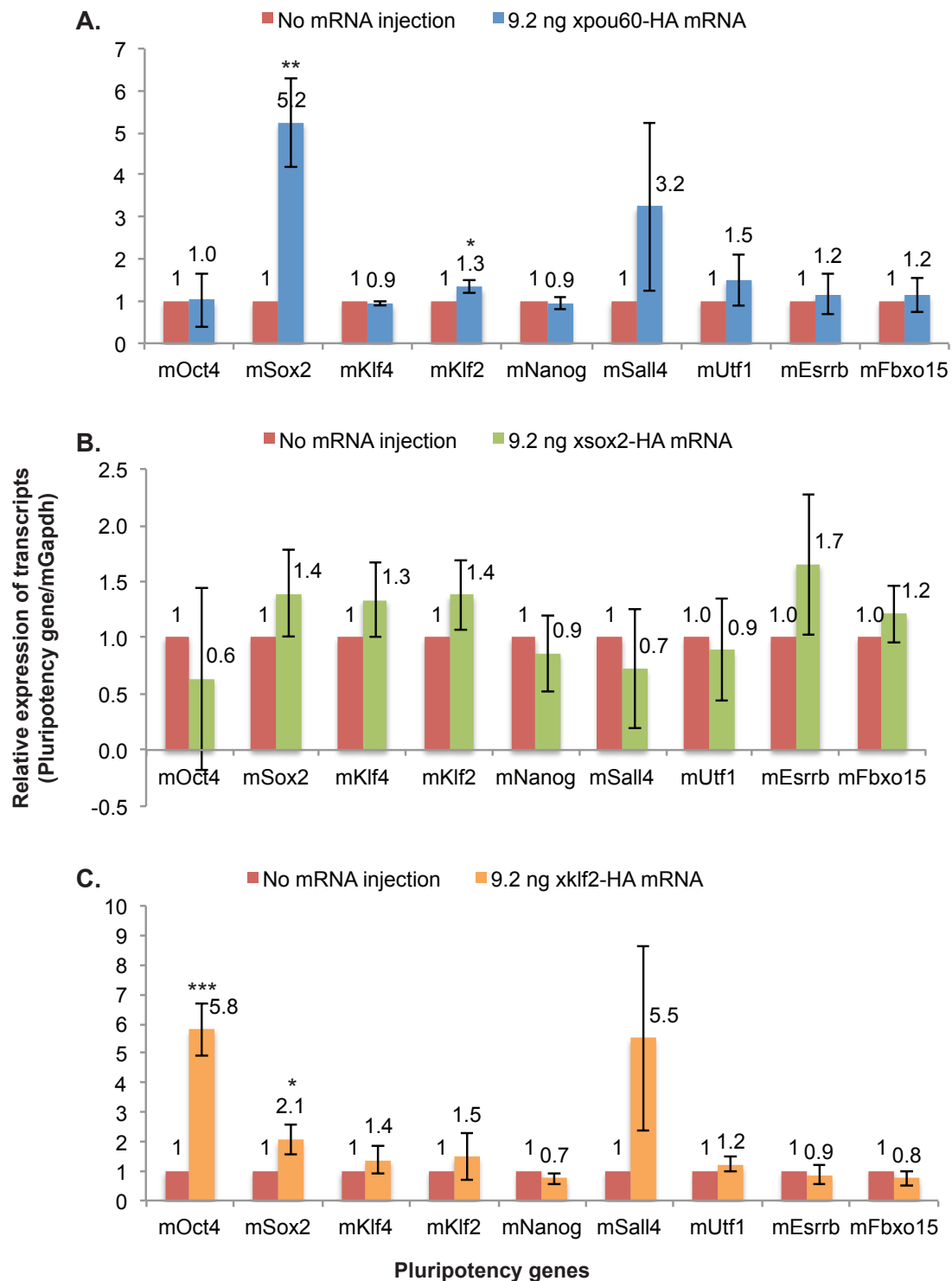


Figure 3.3.3 The overexpression of *Xenopus* Yamanaka homologs up-regulates different sets of selected pluripotency genes in MEFs at Day 2 after Oocyte-NT.

The relative expression of transcripts in control group (No mRNA injection) is 1 for each pluripotency gene and shown as a bar in red. (Figure legend continues on the next page)

(A) The relative expression of transcripts in xpou60-HA overexpression group (9.2 ng xpou60-HA mRNA) for each pluripotency gene represents the effect of the overexpression of xpou60-HA proteins and shown as bars in blue (* $p < 0.05$, ** $p < 0.01$, $n = 3$, t -test).

(B) The relative expression of transcripts in xsox2-HA overexpression group (9.2 ng xsox2-HA) for each pluripotency gene represents the effect of the overexpression of xsox2-HA proteins and shown as bars in green ($n = 3$, t -test).

(C) The relative expression of transcripts in xklf2-HA overexpression group (9.2 ng xklf2-HA mRNA) for each pluripotency gene represents the effect of the overexpression of xklf2-HA proteins and shown as bars in orange (* $p < 0.05$, *** $p < 0.001$, $n = 3$, t -test).

3.3.3 Summary

Overall, protein synthesis and gene regulation of *Xenopus* YF homologs are evaluated in this section. It shows that proteins of *Xenopus* YF homologs are synthesized from injected mRNA dose- and time-dependently in *Xenopus* oocytes. Furthermore, the accumulation of *Xenopus* YF homologs in GV of *Xenopus* oocytes is dependent to the dose of injected mRNA.

Importantly, in this section, it shows that the regulation of pluripotency genes by YF homologs is conserved across species since *Xenopus* YF homologs regulate similar sets of tested pluripotency genes in mammalian nuclei as their mammalian counterparts. While hOCT4-HA up-regulates the expression of mSox2 by 3.6-fold in sixiFM MEFs, xpou60 up-regulates the expression of mSox2 by 5.2-fold and mSall4 by 3.2-fold in the sixiFM MEF nuclei. While hKLF4 up-regulates the expression of mOct4 by 3.3-fold, mSox2 by 4.3-fold, mSall4 by 4-fold and mUtf1 by 2.9-fold in sixiFM MEFs, xklf2-HA up-regulates the expression of mOct4 by 5.8-fold, mSox2 by 2.1-fold and mSall4 by 5.5-fold. Therefore, the *Xenopus* oocyte system can demonstrate the inter-species regulation of genes since mammalian and *Xenopus* YF homologs can regulate similar sets of genes.

However, xsox2-HA, unlike mSox2-HA, cannot up-regulate any of tested pluripotency genes while mSox2-HA can up-regulate mOct4 by 2.6-fold, mSox2 by 14-fold and mKlf4 by 2.3-fold in sixiFM MEFs. This indicates that the different regulation of genes by YF homologs can happen possibly due to the presence of specific cofactors and the amount of these cofactors.

3.4 Conclusions

In Section 3.2.1 and 3.3.1, the production and localization of mammalian and *Xenopus* YF proteins are examined to confirm that these YFs can be translated from injected YF mRNA in the cytosol and transported to the GV where they bind to their target genes. In Section 3.2.2 and 3.3.2, it has proven that the overexpression of each YF can regulate different sets of selected pluripotency genes in the MEFs via Oocyte-NT.

However, since the experimental settings for Oocyte-NT is based on interspecies interaction, it would be rational to understand how the extent of interspecies interaction would affect the regulation of downstream genes by transcription factors. In this chapter, test YFs belong to four species: *Homo sapien* (human, hOCT4-HA and hKLF4), *Mus musculus* (house mouse, mSox2-HA), *Xenopus laevis* (Africa clawed frog, xpou60-HA and xklf2-HA) and *Xenopus tropicalis* (Western clawed frog, xsox2-HA). Additionally, human and house mouse are within the same class, mammal; Africa clawed frog and Western clawed frog are within the same genus, *Xenopus*.

During Oocyte-NT, the recipient cells are oocytes of Africa clawed frog and the donor cells are embryonic fibroblasts of house mouse. As a result, they will share the contents in the cells and there will be crosstalk between their molecules. Based on this point, to evaluate if transcription factors from one species can regulate downstream genes of another species during Oocyte-NT, the interspecies interactions would need to be considered for the data interpretation.

For this matter, I took the data from 3.2.2 and 3.3.2 and made two tables for easy comparison among test YFs and evaluate how conserved it is for mammalian and *Xenopus* YFs to regulate pluripotency genes in mouse genome (Table 3.4). Since the regulation of downstream genes by members of the same TF families is context-dependent¹¹⁵⁻¹¹⁷, it is sensible that homologs of the same YF families regulate different sets of pluripotency genes.

Among these three YF families, KLF family members, hKLF4 and xklf2-HA, upregulated the most of selected pluripotency genes in MEF and it suggests KLF members are able to activate pluripotency genes in the Oocyte-NT system. Therefore, I chose xklf2-HA as the major transcription factor for the following experiments to evaluate the role of transcription factors during SCNR by oocytes.

Table 3.4

	mOct4	mSox2	mKlf4	mKlf2	mNanog	mSall4	mUtf1	mEsrrb	mFbxo15
<i>Mammalian YFs</i>									
hOCT4-HA (n=3)	1.4	3.6	0.9	1.2	1.2	1.3	1	0.9	1.2
mSox2-HA (n=3)	2.6	14.4	2.3	1	1.1	0.7*	1.2	1.8	1.1
hKLF4 (n=2)	3.3*	4.3	1.1	1.3**	1.3	4*	2.9*	1.1	1.2
<i>Xenopus YFs</i>									
xpou60-HA (n=3)	1	5.2**	0.9	1.3*	0.9	3.2	1.5	1.2	1.2
xsox2-HA (n=3)	0.6	1.4	1.3	1.4	0.9	0.7	0.9	1.7	1.2
xklf2-HA (n=3)	5.8***	2.1*	1.4	1.5	0.7	5.5	1.2	0.9	0.8

Table 3.4.A Fold change of pluripotency genes in sixiFM MEFs between Control groups (No mRNA injection) and YF overexpression groups (9.2 ng YF mRNA) at Day 2 after Oocyte-NT is shown.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, *t*-test

	mOct4	mSox2	mKlf4	mKlf2	mNanog	mSall4	mUtf1	mEsrrb	mFbxo15
<i>Mammalian YFs</i>									
hOCT4-HA (n=3)		+							
mSox2-HA (n=3)	+	+	+						
hKLF4 (n=2)	+	+				+	+		
<i>Xenopus YFs</i>									
xpou60-HA (n=3)		+				+			
xsox2-HA (n=3)									
xklf2-HA (n=3)	+	+				+			

Table 3.4.B Up-regulation of pluripotency genes in sixiFM MEFs by the overexpression of mammalian and *Xenopus* YFs at Day 2 after Oocyte-NT is shown.

+ means the up-regulation of pluripotency genes by the overexpression of YFs

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, *t*-test

Chapter 4 Time-point observation demonstrates gene regulation in mammalian nuclei by oocyte factors and overexpression of xklf2-HA

4.1 Introduction

4.1.1 Background

In chapter 3, it has been shown that interspecies regulation of downstream genes is achievable through Oocyte-NT, and that paired Yamanaka factor (YF) homologs can regulate similar sets of pluripotency genes during SCNR by *Xenopus* oocytes. However, the increased relative expression is variable and this may relate to the time points of gene activation. Therefore, I have used time-point observations after Oocyte-NT to examine gene activation by oocyte factors and xklf2-HA overexpression.

Among tested Yamanaka homologs, KLF family members, hKLF4 and xklf2-HA, exert the strongest effect on the selected pluripotency genes in MEFs. Additionally, xklf2 is predicted to be an oocyte factor in *Xenopus* oocytes and eggs based on its mRNA expression profile¹⁴⁰. It has been shown that xklf2 mRNA is expressed before MBT and to decline afterwards and its dynamic mRNA expression profile correlates with the protein expression profile from fertilization to hatching of *Xenopus* embryos¹⁴⁰⁻¹⁴².

Hence, in this chapter, I have chosen *xklf2*-HA to mimic the maternal transcription factors in *Xenopus* oocytes and ask how *xklf2*-HA overexpression affects the regulation of pluripotency genes in MEFs and mMyoblasts (mMyo) during SCNR by oocytes.

4.1.2 Experimental design

To distinguish the effects of *Xenopus* oocyte factors from ectopic *xklf2*-HA, Oocyte-NT is performed with or without *xklf2*-HA overexpression and samples are collected at multiple time points after Oocyte-NT. Comparing the samples at multiple time points from time 0 when performing Oocyte-NT to later time points demonstrates the progressive increase or decrease of transcripts regulated by oocyte factors during SCNR. Comparing the samples collected at the same time point reveals how *xklf2*-HA overexpression affects gene expression during SCNR by oocytes.

In the experimental settings (Figure 4.1), 9.2 ng of *xklf2*-HA mRNA was injected into each *Xenopus* oocyte 24 hours before Oocyte-NT. Three types of cell lines were used for Oocyte-NT. sixiFM and TcR2 are cell lines of mouse embryonic fibroblasts (MEFs) and C2C12 is a cell line of mouse myoblasts (mMyos). Samples were collected directly after performing Oocyte-NT at Day 0 or collected at days between Day 0 and Day 3. The RNA of samples was then extracted and analyzed by QPCR.

The expression of pluripotency genes is normalized to the expression of mGapdh and mGapdh level is also an indication of the number of cells

transplanted into GVs of oocytes. Six of the selected pluripotency genes, namely mOct4, mSox2, mSall4, mEsrrb, mNanog and mFbxo15, are shown to be bound by mKlf2/4/5 on their promoters, enhancers or gene bodies via ChIP analysis¹¹⁸ and all of them have one or two mKlf2 binding motif on their promoters according to TRANSFAC database¹⁴³.

Figure 4.1

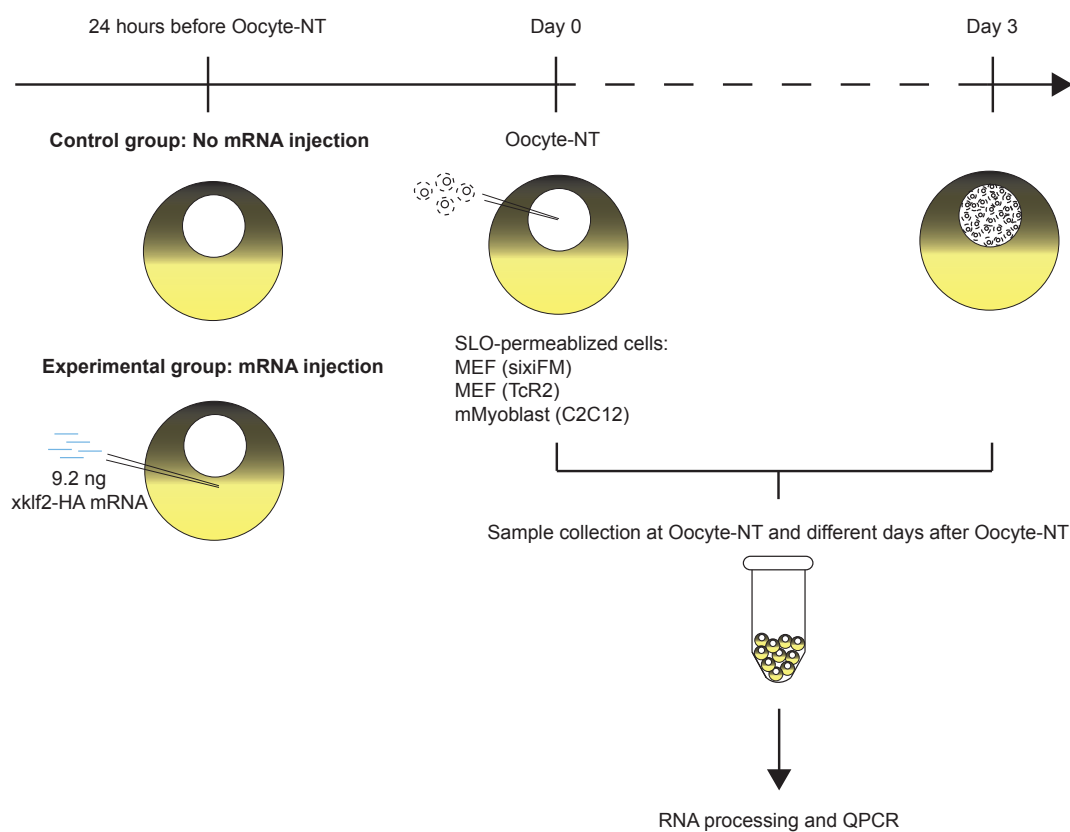


Figure 4.1 Sample preparation for comparing the effects of oocyte factors at multiple time points and xklf2-HA overexpression at fixed time points on the regulation of selected pluripotency genes is shown.

4.2 The overexpression of xklf2-HA regulates some tested pluripotency genes in sixiFM MEFs and may be an oocyte factor during SCNR by *Xenopus* oocytes

To compare the effect of oocyte factors with xklf2-HA overexpression on regulating selected pluripotency genes during SCNR by *Xenopus* oocytes, the samples are collected at Day 0, 1, 2 and 3 after Oocyte-NT with or without xklf2-HA mRNA injection. SLO-permeabilized sixiFM MEFs were used for Oocyte-NT in this section.

4.2.1 Oocyte factors up-regulate mOct4, mSox2, mKlf4, mSall4, mUtf1 and mEsrrb in sixiFM MEFs from Day 0 to Day 3 after Oocyte-NT

Comparing the relative expression of transcripts in the no mRNA injection groups at different time points (red lines, Figure 4.2.1, 4.2.2 and 4.2.3), the trend of the lines for the pluripotency genes shows how oocyte factors affect their expression.

For mOct4 (red line, Figure 4.2.1.A), the relative expression at Day 0 is undetermined by QPCR due to nil or low expression of mOct4 in the sixiFM MEFs. Then oocyte factors up-regulate the expression of mOct4 to a detectable level from Day 0 to Day 1 and by 2.9-fold from Day 1 to Day 3. The up-regulation of mSox2 by oocyte factors is strong from Day 0 to Day 3 by 243-fold (red line, Figure 4.2.1.B) and, comparably, mKlf4, mSall4 and mUtf1 have relatively mild increases in their expression from Day 0 to Day 3 by 12-fold, 33-fold and 19-fold (red lines, Figure 4.2.1.C-D and 4.2.2.A), respectively.

The expression of mEsrrb is up-regulated by oocyte factors gently and steadily with a 4.6-fold increase from Day 0 to Day 3 (red line, Figure 4.2.2.B).

For mKlf2, mNanog and mFbxo15 (Figure 4.2.3), it seems the expression of these genes is negatively regulated by oocyte factors by 0.4-fold, 0.3-fold and 0.4-fold from Day 0 to Day 3. However, they could be unaffected by oocyte factors and degraded due to the post-transcriptional mechanisms in the donor cells or *Xenopus* oocytes.

4.2.2 xklf2-HA overexpression augments the up-regulation of mOct4, mSox2, mKlf4 and mSall4 by oocyte factors by oocyte factors at Day 1, 2 and 3 after Oocyte-NT

Overall, the trend of xklf2-HA mRNA groups shows that xklf2-HA overexpression augments the up-regulation of mOct4, mSox2, mKlf4 and mSall4 by oocyte factors at Day 1, 2 and 3 after Oocyte-NT (Figure 4.2.1).

mUtf1 and mEsrrb have been shown to be up-regulated by oocyte factors but they are not affected by xklf2-HA overexpression (Figure 4.2.2). Furthermore, it is the same for oocyte factors that mKlf2, mNanog and mFbxo15 are not affected by xklf2-HA overexpression (Figure 4.2.3)

4.2.3 The fluctuating augmentation caused by xklf2-HA overexpression explains the stochastic nature of oocyte

When comparing the relative expression of xklf2-HA mRNA groups to the relative expression of no mRNA injection groups at the same time points, the effect of xklf2-HA on the up-regulation of pluripotency genes by oocyte factors can be observed at all time points for mOct4, mSox2, mKlf4 and mSall4 (red and orange dots, Figure 4.2.1). Take mOct4 as an example (Figure 4.2.1.A), xklf2-HA overexpression augments the up-regulation of mOct4 by oocyte factors by 5.5 times at Day 1, by 23 times at Day 2 and by 11 times at Day 3.

However, each pluripotency gene responds to xklf2-HA overexpression at different time points and to a different extent. Among these four genes (Figure 4.2.1), only the up-regulation of mSox2 by oocyte factors is augmented by xklf2-HA overexpression with steadily increasing values by 1.5-fold at Day 0, 1.9-fold at Day 1, 2.8-fold at Day 2 to 3.2-fold at Day 3 (Figure 4.2.1.B). For mOct4 and mSall4 (Figure 4.2.1.A and D), the augmentation by xklf2-HA overexpression is strong with a maximum 22-fold and 20-fold increase respectively, at Day 2 but this augmentation fluctuates throughout this period at each Day after Oocyte-NT. For mKlf4 (Figure 4.2.1.C), this augmentation also fluctuates. This phenomenon can explain why sometimes the effect of transcription factors is obvious but the statistical significance is hard to reach due to the stochastic nature of each oocyte even if all the oocytes are taken from the same ovary of one female frog.

4.2.4 The up-regulation of pluripotency genes by oocyte factors happens most strongly within one day after Oocyte-NT and the one-day regulation is sufficient to represent the long-term regulation from Day 0 to Day 3

Comparing the relative expression of each pluripotency gene between Day 0 and Day 1 after Oocyte-NT, the scale of up-regulation for each gene by oocyte factors is stronger than at other time points within 1-day period (red dots, Figure 4.2.1). For example, the relative expression of mSox2 from Day 0 to Day 3 increases by 243-fold and it is more than the fold change of 33 for mSall4 from Day 0 to Day 3 (red dots, Figure 4.2.1.B and D). Likewise, this scale of up-regulation can be decided by the increase of relative expression from Day 0 to Day 1. For example, the increased fold change for mSox2 is 107 and it is more than 12 for mSall4 (red dots, Figure 4.2.1.B and D).

Additionally, the relative expression increases most obviously from Day 0 to Day 1, compared to other relative expression difference within one day. Take mSox2 as an example (red dots, Figure 4.2.1.B), the increased fold change from Day 0 to Day 1 is 107 ($107/1=107$) and it is 59 times more than the increased fold change from Day 1 to Day 2, which is 1.8 ($=193/107$). Also, it is 82 times more than the increased fold change from Day 2 to Day 3, which is 1.3 ($=243/193$).

These two points suggest that SCNR by oocyte happens most vigorously within one day after Oocyte-NT. This allows further manipulation to be applied after Oocyte-NT while Oocyte-NT experiments usually end within two days and *Xenopus* oocytes tend to deteriorate over time after Oocyte-NT.

4.2.5 Up-regulation of mOct4, mSox2, mKlf4, mSall4 and mUtf1 by xklf2-HA overexpression starts at different time points after Oocyte-NT

It has been shown that mOct4, mSox2, mKlf4 and mSall4 are up-regulated and mUtf1 is unaffected by xklf2-HA overexpression at Day 2 after Oocyte-NT (Figure 4.2.1 and 4.2.2.A). Notably, xklf2-HA overexpression starts to up-regulate the expression of these pluripotency genes at different time points after Oocyte-NT, including mUtf1, which is judged to be unaffected by xKlf2-HA. For mOct4 and mSall4, the expression is up-regulated by xklf2-HA overexpression by 5.5-fold and 3.7-fold at Day 1 after Oocyte-NT (red and orange dots at Day 1, Figure 4.2.1.A and D). In the case of mSox2 and mKlf4, xklf2-HA overexpression starts to up-regulate these two genes by 2.8-fold and 2.3-fold at Day 2 after Oocyte-NT (red and orange dots at Day 2, Figure 4.2.1.B-C). For mUtf1, it is unaffected at Day 1 and Day 3 after Oocyte-NT but it is up-regulated mildly at Day 0 by 2.3-fold and Day 2 by 2-fold after Oocyte-NT by xklf2-HA overexpression (red and orange dots, Figure 4.2.2.A).

Since pluripotency genes usually remained silent in MEFs, the different start points of up-regulation by xklf2-HA overexpression may suggest that pluripotency genes are released from closed chromatin and become accessible by transcription factors at different time points after Oocyte-NT.

4.2.6 xklf2-HA overexpression up-regulates some selected pluripotency genes as oocyte factors and it may be one of the effective oocyte factors

Although all the selected pluripotency genes are direct binding targets of mKlf2 as previously mentioned and some of them indeed are regulated by xklf2-HA overexpression, xklf2-HA selectively regulates the genes, which are also regulated by oocyte factors (Figure 4.2.1 and 4.2.2). There is no doubt that some transcription factors already exist in the oocytes to up-regulate xklf2-HA target gene expression. Therefore, it suggests that xklf2 may be one of the oocyte factors to regulate some xklf2 target genes and that some other oocyte factors may function redundantly as xklf2.

Although overexpressed xklf2-HA has a substantial effect on tested genes here, the overwhelmingly strong conclusion is that this factor overexpression is, in most cases, small by comparison with the effect of endogenous oocyte factors in the same time period.

Figure 4.2.1

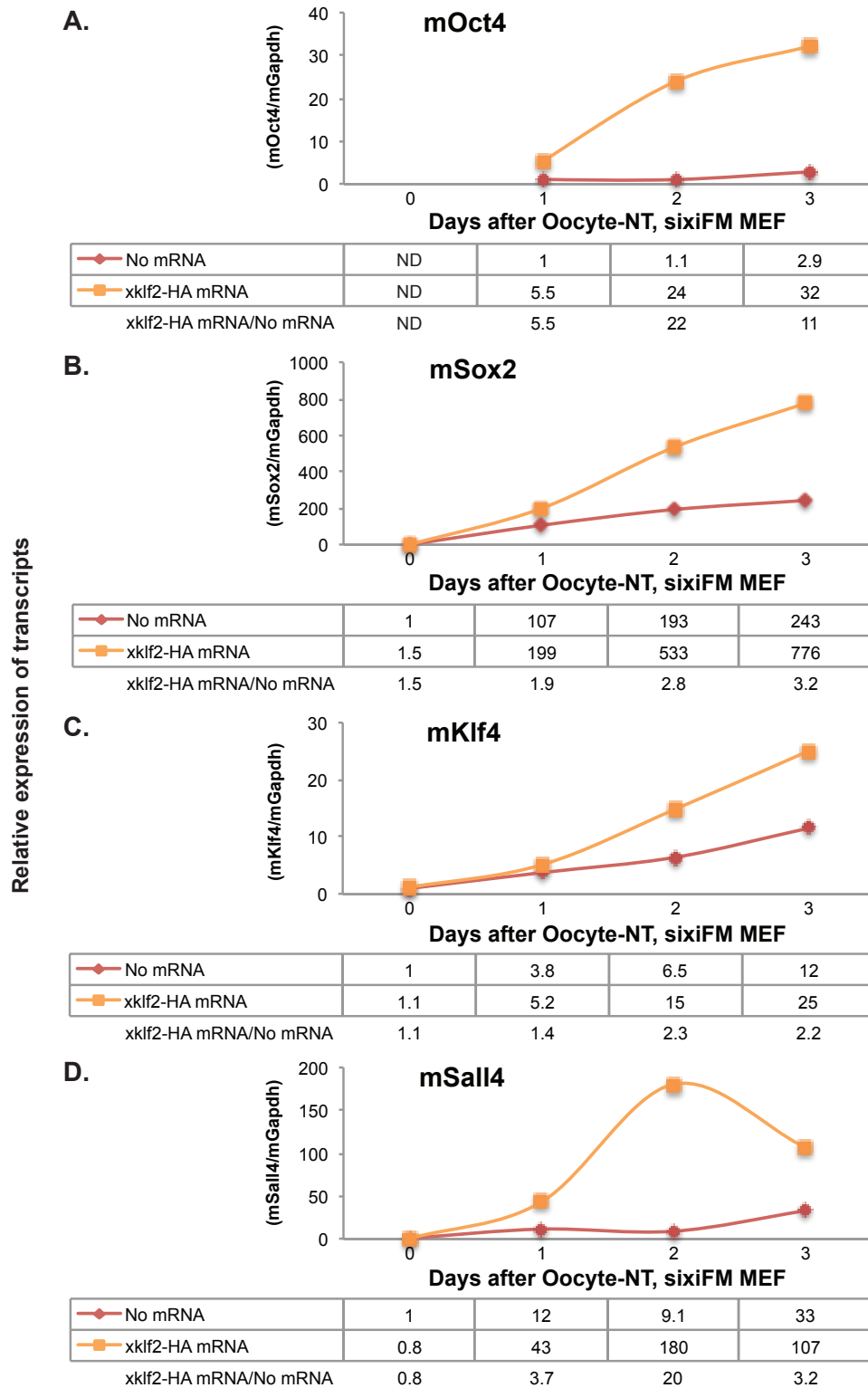


Figure 4.2.1 mOct4, mSox2, mKlf4 and mSall4 in sixiFM MEFs are up-regulated strongly by oocyte factors with fold changes of up to 243 after Oocyte-NT and this up-regulation is augmented by xklf2-HA overexpression with fold change of up to 22.

Figure 4.2.2

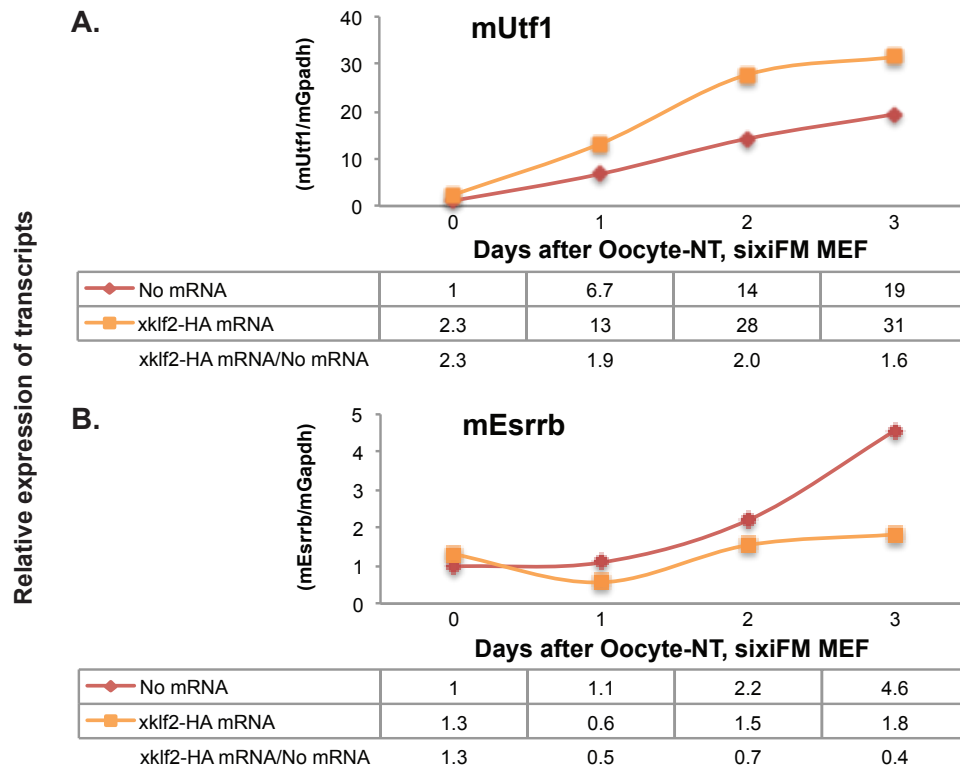


Figure 4.2.2 mUtf1 and mEsrrb in sixiFM MEFs are up-regulated by oocyte factors but un-affected by xklf2-HA overexpression after Oocyte-NT.

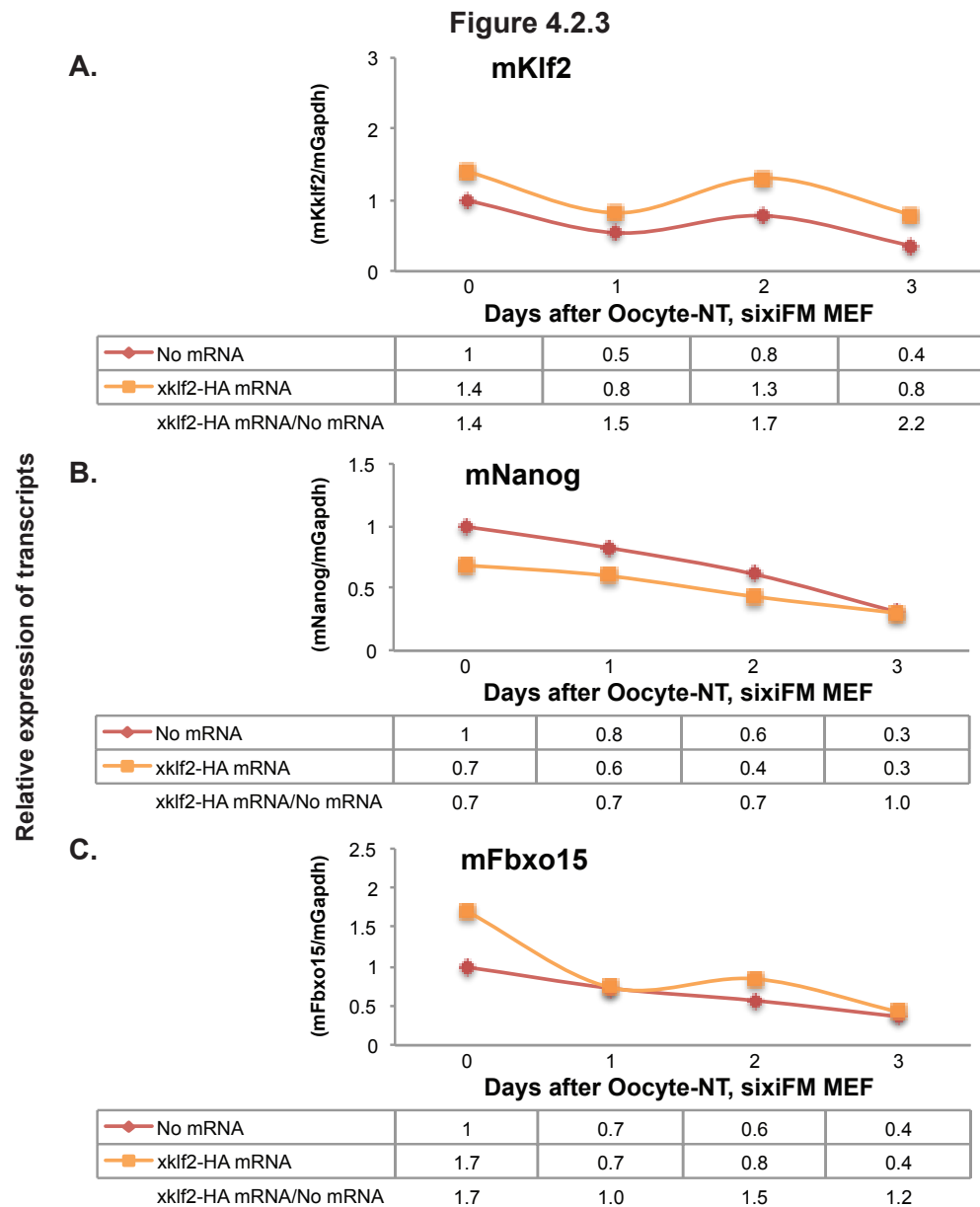


Figure 4.2.3 mKlf2, mNanog and mFbxo15 in sixiFM MEFs are un-affected by oocyte factors and xklf2-HA overexpression after Oocyte-NT.

4.2.7 Summary

To sum up, the effect of overexpressed xklf2-HA on expression of 9 mouse pluripotency genes in transplanted sixiFM MEFs has been tested in *Xenopus* oocytes. Through a time-point observation, 6 of the tested pluripotency genes were up-regulated strongly by oocyte factors by up to 243-fold from Day 0 to Day 3 after Oocyte-NT and 4 of tested pluripotency genes were up-regulated mildly by xklf2-HA overexpression by up to 22-fold at Day 2 after Oocyte-NT.

The up-regulation of pluripotency genes by oocyte factors is most strongly seen within one day after Oocyte-NT and these genes are continuously up-regulated mildly by oocyte factors afterwards. Overexpressed xklf2-HA redundantly up-regulates the genes that are also up-regulated by oocyte factors and is possibly one of the effective transcription factors in *Xenopus* oocytes.

4.3 The overexpression of xklf2-HA facilitates the expression of pluripotency genes in MEFs at the beginning of SCNR by *Xenopus* oocytes

To investigate the timing of enhanced gene expression by xklf2-HA overexpression, the expression of pluripotency genes is compared in different ways and another MEF cell line, TcR2, is used for Oocyte-NT in this section. The effect of oocyte factors on regulation of pluripotency genes is evaluated by comparing gene expression in no mRNA injection groups at Day 0 with gene expression at Day 2 after Oocyte-NT (Figure 4.3.A). The effect of xklf2-HA overexpression is evaluated at Day 0 and Day 2 after Oocyte-NT in order to demonstrate the immediate and long-term influence of xklf2-HA on regulation of pluripotency genes during SCNR by *Xenopus* oocytes (Figure 4.3.B and C).

4.3.1 Up-regulation of pluripotency genes in TcR2 MEFs by oocyte factors and xklf2-HA overexpression is similar to the up-regulation of genes in sixiFM MEFs

Similar to sixiFM MEFs, the expression of mOct4, mSox2 and mUtf1 in TcR2 MEFs are up-regulated strongly by oocyte factors from Day 0 to Day 2 after Oocyte-NT by 92-fold, 39-fold and 13-fold, respectively (Figure 4.3.A). Additionally, the expression of mJun in TcR2 MEFs is also up-regulated by oocyte factors by 7.3-fold (Figure 4.3.A).

Likewise, the expression of mOct4 and mSox2 in TcR2 MEFs is up-regulated mildly by xklf2-HA overexpression at Day 2 after Oocyte-NT (Figure 4.3.C). At Day 2 after Oocyte-NT, mOct4 and mSox2 are up-regulated by xklf2-HA overexpression by 3.4-fold and 6.4-fold, respectively (Figure 4.3.C). Interestingly, mOct4 and mUtf1 are up-regulated by xklf2-HA overexpression at Day 0 after Oocyte-NT with fold changes of 41-fold and 3.7-fold (Figure 4.3.B), which are much bigger than the fold changes of the same genes up-regulated by xklf2-HA overexpression at Day 2 after Oocyte-NT (Figure 4.3.C).

4.3.2 Early and late response of pluripotency genes to xklf2-HA overexpression at Day 0 and Day 2 after Oocyte-NT

It has been shown that pluripotency genes in sixiFM MEFs are regulated by xklf2-HA overexpression at different time points after Oocyte-NT (Figure 4.2.1, page 110). Here, it also has been shown that mOct4 and mUtf1 in TcR2 MEFs are up-regulated by xklf2-HA overexpression at Day 0 after Oocyte-NT (Figure 4.3.B) and mOct4 and mSox2 are up-regulated at Day 2 after Oocyte-NT (Figure 4.3.C).

It seems the effects of xklf2-HA overexpression on different genes starts at different time points and genes affected by xklf2-HA overexpression can be categorized into two responses. mOct4 and mUtf1 are categorized as the early response genes and mSox2 is categorized as late response genes. Take mOct4 as an example, the fold change between xklf2-HA mRNA groups and no mRNA injection groups at Day 0 is 41-fold and it is more than the fold change of 3.7 at Day 2 after Oocyte-NT (Figure 4.3.B-C). On the contrary, the

up-regulation of mSox2 by xklf2-HA overexpression at Day 0 is smaller than the up-regulation of mSox2 by xklf2-HA overexpression at Day 2 while the fold changes are 1.6 at Day 0 and 6.4 for Day 2 after Oocyte-NT (Figure 4.3.B-C).

The difference in the starting time points of xklf2-HA effects suggests that the regulation process by xklf2-HA overexpression may be affected by other mechanisms. A possible one is the original chromatin structures of genes in donor cells and the accessibility of chromatin structures may hinder or delay the binding of transcription factors, and hence, the gene activation.

4.3.3 mOct4 is a SCNR resistant gene in MEFs and is up-regulated by xklf2-HA overexpression immediately at Day 0 after Oocyte-NT with the highest fold change among all test pluripotency genes

In section 4.2, the expression of mOct4 is not determined in the sixiFM MEFs (Figure 4.2.1.A, page 110) while the expression of mOct4 is detectable in the TcR2 MEFs due to the increase of reversed transcribed cDNA for QPCR analysis (Figure 4.3). From a nil or low expression state in MEFs, mOct4 was activated to a detectable level by oocyte factors one day after Oocyte-NT (Figure 4.2.1.A, page 110) and it might suggest mOct4 is an resistant gene to SCNR by oocytes. Interestingly, the expression of mOct4 in TcR2 MEFs is up-regulated by the oocyte factors from Day 0 to Day 2 after Oocyte-NT by 92-fold (Figure 4.3.A) and by xklf2-HA overexpression at Day 0 by 41-fold (Figure 4.3.B) with the highest fold change among all test pluripotency genes.

This shows that the nil or lowly expressed genes in MEFs, such as mOct4, would be up-regulated by oocyte factors more strongly than other middle or highly expressed genes, such as mSox2, mUtf1 and mJun (Figure 4.3.A). Additionally, the early response of mOct4 to xklf2-HA overexpression at Day 0 after Oocyte-NT also suggests the ability of xklf2-HA overexpressoin to access the closed chromatin and activate downstream genes (Figure 4.3.B).

Figure 4.3

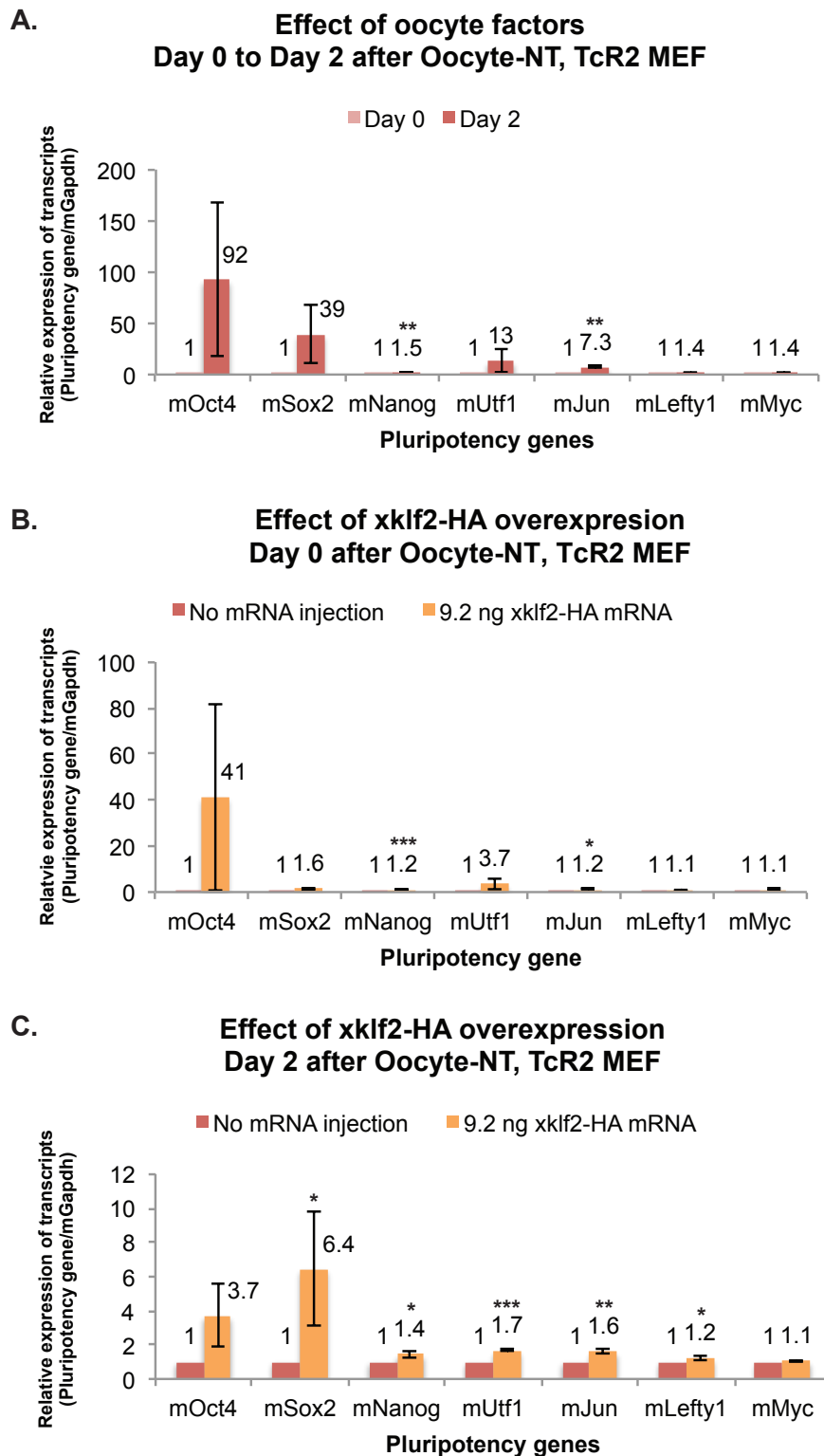


Figure 4.3 The immediate and long-term effect of xklf2-HA overexpression on regulation of pluripotency genes in TcR2 MEFs.

(Figure legend continues on the next page)

mOct4, mSox2, mNanog, mUtf1 and mJun are up-regulated by both maternal factors from Day 0 to Day 2 and the overexpression of xklf2-HA at either Day 0 or Day 2 after Oocyte-NT. mOct4, mSox2, mNanog (n=3, *t*-test); mUtf1, mJun, mLefty1, mMyc (n=2, *t*-test). **p*<0.05, ***p*<0.01, ****p*<0.001

(A) The effect of oocyte factors on pluripotency genes from Day 0 to Day 2 after Oocyte-NT

(B) The immediate effect of xklf2-HA on pluripotency genes at Day 0 of Oocyte-NT. In all these figures, ng means the amount of xklf2-HA mRNA. Effect at Day 0 is an early response of genes to xklf2-HA overexpression and samples may incubate in oocytes for maximal 10 mins because it takes about 10 mins to finish nuclear transfer for 10 oocytes.

(C) The long-term effect of xklf2-HA on pluripotency genes at Day 2 after Oocyte-NT

4.3.4 Summary

All in all, oocyte factors and xklf2-HA overexpression up-regulate the same sets of pluripotency genes in TcR2 and sixiFM MEFs. Therefore, chromatin structures may be similar between TcR2 and sixiFM MEFs. In addition, the SCNR resistant gene, mOct4, is up-regulated most strongly by oocyte factors among tested pluripotency genes.

Furthermore, the immediate response of mOct4 to xklf2-HA overexpression at Day 0 after Oocyte-NT suggest the ability of xklf2-HA overexpression to activate downstream genes when these genes are still resistant to be activated by oocyte factors. As with the genes in sixiFM MEFs, the endogenous factors in oocytes are much more effective than overexpressed xklf2-HA on nearly all tested genes.

4.4 The similar and different response of pluripotency genes in MEFs and mMyoblasts to oocyte factors and to xklf2-HA overexpression

In section 4.2 and 4.3, the regulation of tested pluripotency genes in sixiFM and TcR2 MEFs by oocyte factors and xklf2-HA overexpression is discussed. For both cell lines of MEFs, sixiFM and TcR2, the sets of pluripotency genes regulated by the oocyte factors and xklf2-HA overexpression are probably similar due to their similar cellular characteristics, namely chromatin state and expression of genes. In order to examine the regulation of the same genes in different cell types by oocyte factors and xklf2-HA overexpression, MEFs (TcR2) and mMyos (C2C12) are compared in this section.

4.4.1 Similar response of tested pluripotency genes in MEFs and mMyos to oocyte factors and xklf2-HA overexpression except for the response of mUtf1 in mMyos to oocyte factors

Comparing the Day 2 samples to the Day 0 samples for each gene in MEFs without mRNA injection, oocyte factors strongly up-regulate the expression of mOct4 by 27-fold, mSox2 by 41-fold, mUtf1 by 5.6-fold and mJun by 6.3-fold (solid red lines, Figure 4.4.1 and 4.4.2). Similar to the effect of oocyte factors on MEFs, mOct4, mSox2 and mJun in mMyos are also up-regulated by oocyte factors by 437-fold, 577-fold and 4-fold from Day 0 to Day 2 after Oocyte-NT (dotted red lines, Figure 4.4.1 and 4.4.2.B) except that mUtf1 is un-affected (dotted red lines, Figure 4.4.2.A). For mNanog, mMyc and

mLefty1, these genes in both MEFs and mMyos are unaffected by oocyte factors or xklf2-HA overexpression (Figure 4.4.3).

The similar sets of up-regulated genes in MEFs and mMyos by oocyte factors except for mUtf1 in mMyos suggest the difference of intrinsic nature between MEFs and mMyos. It may also suggest the mUtf1 in mMyo is resistant to oocyte factors while mUtf1 in MEF can be up-regulated by these factors.

On top of the effect of oocyte factors, xklf2-HA overexpression enhances the up-regulation of mOct4 and mSox2 in MEFs by oocyte factors at Day 0 and Day 2 after Oocyte-NT (solid red and orange dots, Figure 4.4.1). At Day 0, mOct4 in MEFs is up-regulated by 13-fold and mSox2 in MEFs is unaffected by xklf2-HA overexpression (Figure 4.4.1). At Day 2 after Oocyte-NT, both mOct4 and mSox2 in MEFs are up-regulated by xklf2-HA overexpression by 5.3-fold and 2.8-fold, respectively (Figure 4.4.1). In mMyos, mOct4 is up-regulated by xklf2-HA overexpression at Day 0 after Oocyte-NT but is unaffected at Day 2 after Oocyte-NT (Figure 4.4.1.A). Additionally, mSox2 in mMyos is up-regulated by xklf2-HA overexpression at Day 2 but is unaffected at Day 0 after Oocyte-NT (Figure 4.4.1.B).

Therefore, mOct4 in both MEFs and mMyos starts to be up-regulated strongly by xklf2-HA overexpression by 13-fold and 109-fold at Day 0 after Oocyte-NT while mSox2 in both cell types starts to be up-regulated mildly by xklf2-HA overexpression by 2.8-fold and 7.2-fold at Day 2 after Oocyte-NT (Figure 4.4.1). It is consistent to the data shown in section 4.3 that mOct4 in MEFs is

strongly up-regulated by xklf2-HA overexpression by 41-fold at Day 0 and the up-regulation decrease to 3.7-fold at Day 2 after Oocyte-NT (Figure 4.3.B and 4.3.C, page 118). Furthermore, mSox2 in MEFs is not up-regulated by xklf2-HA overexpression at Day 0 but is up-regulated by xklf2-HA overexpression by 6.4-fold at Day 2 after Oocyte-NT (Figure 4.3.B and 4.3.C, page 118). Therefore, the time-dependent response of mOct4 and mSox2 in MEFs to xklf2-HA overexpression is the same as the response of mOct4 and mSox2 in mMyos.

4.4.2 mOct4 and mSox2 in mMyos are up-regulated by oocyte factors and xklf2-HA overexpression more strongly than mOct4 and mSox2 in MEFs

Among all tested pluripotency genes, mOct4 and mSox2 are the most strongly up-regulated by oocyte factors and by xklf2-HA overexpression in both cell types (Figure 4.4.1). Intriguingly, mOct4 and mSox2 respond to oocyte factors and xklf2-HA overexpression differently between MEFs and mMyoblasts.

Comparing the expression of genes increased by oocyte factors, mOct4 in mMyos is up-regulated by oocyte factors by 437-fold, which is more than the fold change of 27 for mOct4 in MEFs up-regulated by oocyte factors from Day 0 to Day 2 after Oocyte-NT (solid and dotted red lines, Figure 4.4.1.A). Likewise, mSox2 in mMyos is up-regulated by oocyte factors by 577-fold and it is more than the fold change of 41 for mSox2 in MEFs up-regulated by oocyte factors from Day 0 to Day 2 after Oocyte-NT (solid and dotted red lines, Figure 4.4.1.B). Hence, the different chromatin states of mSox2 and mOct4 in

MEFs and mMyos probably lead to the different up-regulation of these genes by oocyte factors.

For MEFs, *xklf2*-HA overexpression augments the up-regulation of *mOct4* by oocyte factors by 13-fold at Day 0 and by 5.3-fold at Day 2 after Oocyte-NT (red and orange dots on the solid line, Figure 4.4.1.A). As for mMyos, *xklf2*-HA overexpression augments the up-regulation of *mOct4* by oocyte factors by 109-fold at Day 0 and by 1.5-fold at Day 2 after Oocyte-NT (red and orange dots on the dotted line, Figure 4.4.1.A). Therefore, *mOct4* in mMyos responds to *xklf2*-HA overexpression more strongly than *mOct4* in MEFs at Day 0 after Oocyte-NT (red and orange dots on the dotted line, Figure 4.4.1.A). Additionally, *mSox2* in mMyo is up-regulated at Day 2 after Oocyte-NT more strongly by *xklf2*-HA overexpression by 7.2-fold, compared to *mSox2* in MEF up-regulated by *xklf2*-HA overexpression by 2.8-fold.

Overall, the response of the same genes in different cell types to oocyte factors is the same as the response of these genes to *xklf2*-HA overexpression. That is, the more fold change of gene expression increased by oocyte factors in certain cell types, the more fold change of gene expression increased by *xklf2*-HA overexpression. Therefore, the chromatin structures of genes in different cell types probably affect the response of genes to oocyte factors and *xklf2*-HA overexpression in a similar way.

4.4.3 The nil or low expressed genes in donor cells are more susceptible to the up-regulation induced by oocyte factors and xklf2-HA overexpression

In section 4.3, it suggests that the nil or low expressed genes, such as mOct4 in MEFs, are more sensitive to oocyte factors and xklf2-HA overexpression and up-regulated by oocyte factors by more fold change than other genes (Figure 4.3.A-B, page 118). Based on this point, because mOct4 in mMyos are up-regulated more strongly by oocyte factors (437-fold) and by xklf2-HA overexpression (109-fold) with more fold change than mOct4 in MEFs, which is up-regulated by oocyte factors by 27-fold and by xklf2-HA overexpression by 13-fold, the chromatin structure of mOct4 in mMyos may be more compacted than mOct4 in MEFs (Figure 4.4.1.A).

From the same point, the chromatin structure of mSox2 in mMyos is more compacted than the chromatin structure of mSox2 in MEF because mSox2 in mMyo is more strongly up-regulated by oocyte factors and xklf2-HA overexpression than mSox2 in MEFs (Figure 4.4.1.B).

Therefore, I conclude that the more fold change for genes in certain cell types induced by oocyte factors or xklf2-HA overexpression, the more compacted the chromatin structures of those genes are. That is, the more those genes are resistant to oocyte factors and xklf2-HA overexpression.

Figure 4.4.1

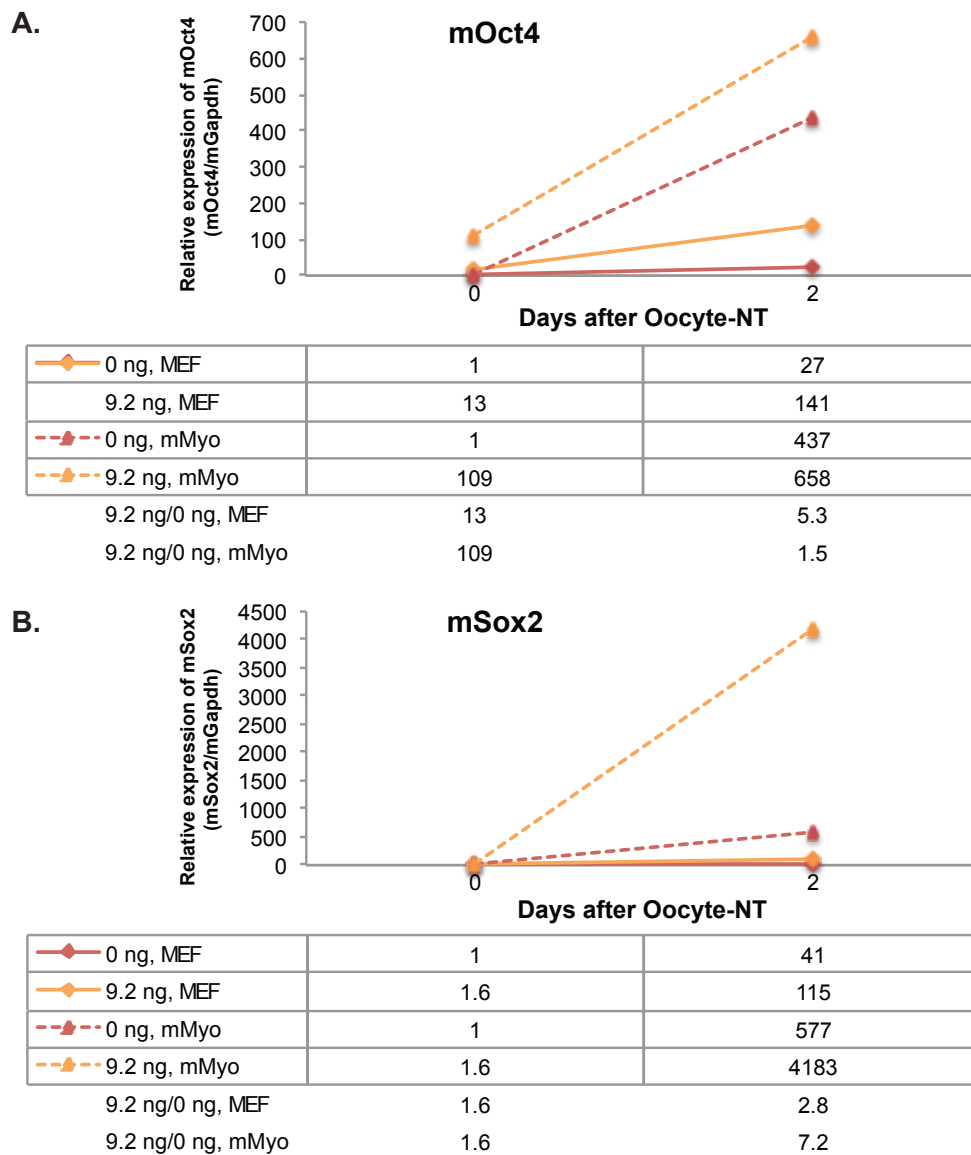


Figure 4.4.1 mOct4 and mSox2 in mMyos are up-regulated more strongly by oocyte factors and xklf2-HA overexpression than mOct4 and mSox2 in MEFs.

In all these figures, ng means the amount of xklf2-HA mRNA. Effect at Day 0 is an early response of genes to xklf2-HA overexpression and samples may incubate in oocytes for maximal 10 mins because it takes about 10 mins to finish nuclear transfer for 10 oocytes.

Figure 4.4.2

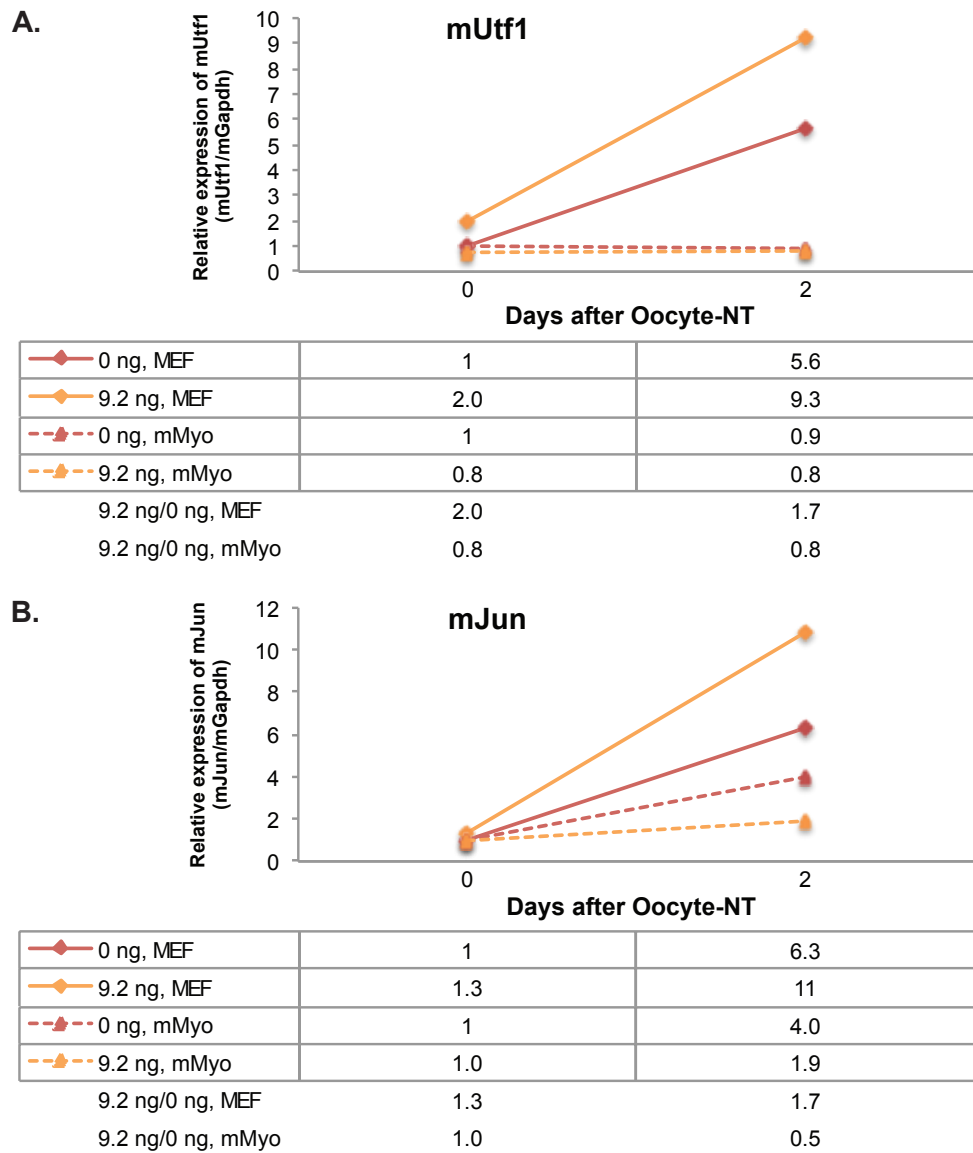


Figure 4.4.2 mJun in both MEFs and mMyos is up-regulated mildly by oocyte factors while mUtf1 in MEFs is up-regulated mildly by oocyte factors but mUtf1 in mMyos is unaffected by oocyte factors.

In all these figures, ng means the amount of xklf2-HA mRNA.

Figure 4.4.3

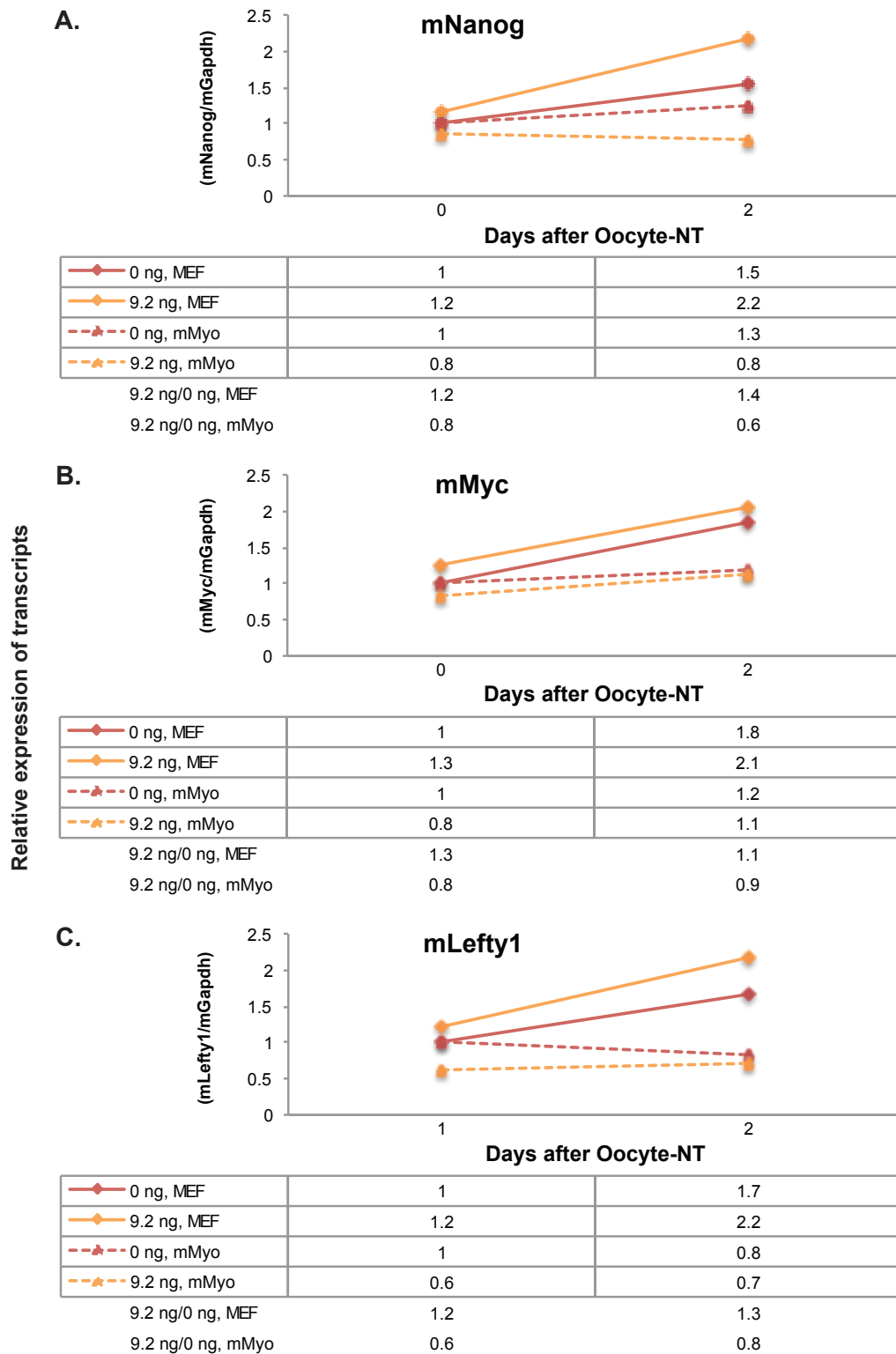


Figure 4.4.3 mNanog, mMyc and mLefty are unaffected by neither oocyte factors nor xklf2-HA overexpression.

In all these figures, ng means the amount of xklf2-HA mRNA.

4.4.4 Summary

To sum, similar sets of tested pluripotency genes in MEFs and mMyos are up-regulated by oocyte factors and xklf2-HA overexpression except for the mUtf1 in mMyos. In addition to effects of oocyte factors on gene regulation, xklf2-HA overexpression regulates downstream pluripotency genes synergistically with oocyte factors and this regulation is cell-type dependent. When genes are strongly up-regulated by oocyte factors in certain cell types, these genes also respond to xklf2-HA overexpression strongly in those cell types. Additionally, this synergistic regulation by xklf2-HA overexpression is time-dependent. For example, mOct4 in both MEFs and mMyos is up-regulated immediately at Day 0 after Oocyte-NT and mSox2 in MEFs and mMyos is up-regulated at Day 2 after Oocyte-NT. Therefore, the up-regulation by oocyte factors and xklf2-HA overexpression is cell-type dependent in terms of the sets of up-regulated genes and the fold change in gene expression and it is due to the different chromatin structures of genes in different cell types.

4.5 Conclusions

In this chapter, I have compared the effects of oocyte factors and *xklf2*-HA overexpression on regulating selected pluripotency genes at different time points after Oocyte-NT. All the tested genes have one or two mKlf2 binding motifs on their promoters¹⁴³. Three cell lines have been used for evaluating the different responses of the same gene in different cell lines to oocyte factors and to *xklf2*-HA overexpression. These cell lines include sixiFM and TcR2 for MEFs and C2C12 for mMyoblasts.

To summarize the effect of oocyte factors on the pluripotency genes in each cell line, the effects are divided into groups according to the fold change (FC) of transcripts between Day 0 and Day 2 samples without *xklf2*-HA mRNA injection (Table 4.5.1). For the down-regulated/unaffected/weakly up-regulated genes ($FC < 2$), oocyte factors do not change much of the expression of mLefty1, mMyc and mNanog in MEFs and mMyoblasts. mJun is mildly up-regulated by oocyte factors in MEFs and mMyoblasts ($2 \leq FC < 16$), and mOct4 and mSox2 are strongly up-regulated in both cell types ($FC \geq 16$).

Notably, mUtf1 in mMyoblasts is down-regulated/unaffected/weakly up-regulated ($FC < 2$) but mUtf1 in MEFs is mildly up-regulated ($2 \leq FC < 16$). Additionally, all the tested pluripotency genes were up-regulated by oocyte factors to the same range of relative expression in both MEF cell lines except for mOct4 while the expression of mOct4 at Day 0 is undetermined in sixiFM MEF so the relative expression is unknown. This indicates the chromatin

structures of tested pluripotency genes in sixiFM and TcR2 MEFs are fairly similar and therefore they respond to oocyte factors in a similar way.

Regarding the effect of *xklf2*-HA overexpression, *xklf2*-HA effect is determined by dividing the relative expression of *xklf2*-HA mRNA injection groups by no mRNA injection groups. There are two different phases for the response of genes to *xklf2*-HA overexpression. The early phase is measured immediately at Day 0 and the late phase is measured at Day 2 after Oocyte-NT.

For the early phase of *xklf2*-HA effect at Day 0 after Oocyte-NT, *mLefty1*, *mMyc*, *mNanog*, *mJun* and *mSox2* are down-regulated/unaffected/weakly up-regulated by *xklf2*-HA overexpression in both MEFs and mMyoblasts ($9.2\text{ng}/0\text{ng}<2$, Table 4.5.2.A). *mUtf1* is down-regulated/unaffected/weakly up-regulated in mMyoblasts ($9.2\text{ng}/0\text{ng}<2$) but is mildly up-regulated in MEFs ($2\leq 9.2\text{ng}/0\text{ng}<16$). *mOct4* is strongly up-regulated in mMEFs and Myoblasts ($9.2\text{ng}/0\text{ng}\geq 16$).

For the late phase of *xklf2*-HA effect at Day 2 after Oocyte-NT (Table 4.5.2.B), *mLefty1*, *mMyc* and *mNanog*, *mUtf1* and *mJun* are down-regulated/unaffected/weakly up-regulated by *xklf2*-HA overexpression in MEFs and mMyoblasts ($9.2\text{ng}/0\text{ng}<2$). *mOct4* is down-regulated/unaffected/weakly up-regulated in mMyoblasts ($9.2\text{ng}/0\text{ng}<2$) while it is mildly up-regulated in MEFs ($2\leq 9.2\text{ng}/0\text{ng}<16$). *mSox2* is mildly up-regulated in both MEFs and mMyoblasts ($2\leq 9.2\text{ng}/0\text{ng}<16$) and no tested gene is strongly up-regulated in both cell types ($9.2\text{ng}/0\text{ng}\geq 16$).

Overall, mLefty1, mMyc and mNanog are down-regulated/unaffected/weakly up-regulated by oocyte factors and by xklf2-HA overexpression in both MEFs and mMyoblasts. mUtf1 is down-regulated/unaffected/weakly up-regulated by oocyte factors and xklf2-HA overexpression in mMyoblasts but it is mildly up-regulated by oocyte factors and xklf2-HA overexpression in MEFs. mOct4 and mSox2 are strongly up-regulated by oocyte factors in both MEFs and mMyoblasts. mOct4 is strongly up-regulated at Day 0 and mSox2 is mildly up-regulated at Day 2 after Oocyte-NT by xklf2-HA overexpression in both MEFs and mMyoblasts.

The similar and different response of tested pluripotency genes in MEFs and mMyos to oocyte factors and xklf2-HA overexpression indicates the way by which oocyte factors and xklf2-HA overexpression regulated gene expression in different cell types due the different chromatin structures of these genes from different donor cell types.

There are only two ways can be utilized to rejuvenate the totipotent or pluripotent state in the cells from adults: one is nuclear transfer and the other one is induced pluripotency. From this point, I will interpret the data acquired here in Chapter 8 Discussion.

Table 4.5.1

	<u>Down-regulated/unaffected/weakly up-regulated (FC<2)</u>						
	<i>mLefty1</i>	<i>mMyc</i>	<i>mNanog</i>	<i>mUtf1</i>	<i>mJun</i>	<i>mOct4</i>	<i>mSox2</i>
MEFs	v	v	v				
mMyoblasts	v	v	v	v			
	<u>Mildly up-regulated (2≤FC<16)</u>						
	<i>mLefty1</i>	<i>mMyc</i>	<i>mNanog</i>	<i>mUtf1</i>	<i>mJun</i>	<i>mOct4</i>	<i>mSox2</i>
MEFs				v	v		
mMyoblasts					v		
	<u>Strongly up-regulated (FC≥16)</u>						
	<i>mLefty1</i>	<i>mMyc</i>	<i>mNanog</i>	<i>mUtf1</i>	<i>mJun</i>	<i>mOct4</i>	<i>mSox2</i>
MEFs						v	v
mMyoblasts						v	v

Table 4.5.1 Effect of endogenous oocyte factors on the regulation of pluripotency genes from Day 0 to Day 2 after Oocyte-NT is shown.

V means the specified genes belong to the groups of down-regulated/unaffected/weakly up-regulated genes (FC<2), mildly up-regulated genes (2≤FC<16) or strongly up-regulated genes (FC≥16).

FC, the fold change of relative expression of genes from Day 0 to Day 2 after Oocyte-NT

Table 4.5.2

<u>Down-regulated/unaffected/weakly up-regulated (9.2ng/0ng<2)</u>							
	<i>mLefty1</i>	<i>mMyc</i>	<i>mNanog</i>	<i>mUtf1</i>	<i>mJun</i>	<i>mOct4</i>	<i>mSox2</i>
MEFs	v	v	v		v		v
mMyoblasts	v	v	v	v	v		v
<u>Mildly up-regulated (2≤9.2ng/0ng<16)</u>							
	<i>mLefty1</i>	<i>mMyc</i>	<i>mNanog</i>	<i>mUtf1</i>	<i>mJun</i>	<i>mOct4</i>	<i>mSox2</i>
MEFs				v			
mMyoblasts							
<u>Strongly up-regulated (9.2ng/0ng≥16)</u>							
	<i>mLefty1</i>	<i>mMyc</i>	<i>mNanog</i>	<i>mUtf1</i>	<i>mJun</i>	<i>mOct4</i>	<i>mSox2</i>
MEFs						v	
mMyoblasts						v	

Table 4.5.2.A The *early* phase of xklf2-HA effect on the regulation of pluripotency genes at Day 0 after Oocyte-NT

9.2ng/0ng=(relative expression of xklf2-HA mRNA samples at Day 0)/(relative expression of no mRNA injection samples at Day 0)

<u>Down-regulated/unaffected/weakly up-regulated (9.2ng/0ng<2)</u>							
	<i>mLefty1</i>	<i>mMyc</i>	<i>mNanog</i>	<i>mUtf1</i>	<i>mJun</i>	<i>mOct4</i>	<i>mSox2</i>
MEFs	v	v	v	v	v		
mMyoblasts	v	v	v	v	v	v	
<u>Mildly up-regulated (2≤9.2ng/0ng<16)</u>							
	<i>mLefty1</i>	<i>mMyc</i>	<i>mNanog</i>	<i>mUtf1</i>	<i>mJun</i>	<i>mOct4</i>	<i>mSox2</i>
MEFs						v	v
mMyoblasts							v
<u>Strongly up-regulated (9.2ng/0ng≥16)</u>							
	<i>mLefty1</i>	<i>mMyc</i>	<i>mNanog</i>	<i>mUtf1</i>	<i>mJun</i>	<i>mOct4</i>	<i>mSox2</i>
MEFs							
mMyoblasts							

Table 4.5.2.B The *late* phase of xklf2-HA effect on the regulation of pluripotency genes at Day 2 after Oocyte-NT

9.2ng/0ng=(relative expression of xklf2-HA mRNA samples at Day 0)/(relative expression of no mRNA injection samples at Day 0)

Chapter 5 Oocytes reprogram gene expression of transcriptomes of mESCs, MEFs and mMyos to an oocyte-steady state except for cell-type specific genes

5.1 Introduction

5.1.1 Background

Mature eggs are able to reprogram somatic cell nuclei into a totipotent state and the resulting totipotent cell can develop into an adult⁹. It has been shown that various somatic cell types can be reprogrammed into a totipotent state in many species but the success rate is usually low due to the resistance of somatic nuclei to SCNR². Therefore, my question is why can oocytes reprogram various cell types? What is the difference or similarity among various cell types against SCNR by oocytes? Are there certain functions of oocytes to reprogram cell nuclei regardless of cell types?

In Chapter 4, it has been shown that the effect of oocyte factors on selected pluripotency genes can be examined by time-point observation via QPCR analysis and the increase of fold change for those genes regulated by oocyte factors is presumably related to the donor cell types due to the different chromatin states and transcriptional machinery of various donor cells. In this chapter, I used RNA-seq plus BrUTP pulldown to examine the transcriptional reprogramming in various cell types by oocyte factors genome wide.

In section 5.2, the time-dependent effect of oocyte factors on MEFs was first compared at Day 1 with Day 2 after Oocyte-NT. I expect to categorize the newly synthesized transcripts of reprogrammed transcriptomes of MEFs into those from up-regulated genes, down-regulated genes and constant genes.

In section 5.3, reprogrammed transcriptomes of mESCs, MEFs and mMyos are compared to find the difference and similarity after various cell types are reprogrammed by oocytes. I expect to find some differentially expressed genes between cell types (Cell-type-DE genes).

In section 5.4, the transcriptomes of each cell type before and after Oocyte-NT will be compared. I expect to have lists of oocyte-off genes and oocyte-on genes. The oocyte-off genes can be either silenced by oocyte factors or not expressed after Oocyte-NT and the oocyte-on genes are activated by oocyte factors after Oocyte-NT.

In section 5.5, the difference and similarity of reprogrammed transcriptomes among various cell types will be compared to understand the ways by which oocytes reprogram various cell types. I expect to have lists of cell-type specific genes and reprogrammable genes among cell types. These cell-type specific genes are either expressed in some cell types but not in other cell types or are resistant to activation by oocyte factors in some cell types but are able to be activated by oocyte factors in other cell types.

5.1.2 Experimental design

To evaluate the effect of oocyte factors on various cell types, Oocyte-NT is followed by BrUTP injection and new synthesized transcripts are identified by BrUTP pull-down before RNA-seq (Figure 5.1.1). This procedure for analyzing induced expression of genes by oocyte factors is similar to experiments in Chapter 4. This is to say that the injection of BrUTP soon after Oocyte-NT marks only transcripts from genes that are newly transcribed at this point and afterwards, rather than paying attention to maternal transcripts already present in oocytes and transplanted cells.

In contrast to SCNR by induced pluripotency, oocytes induce a particularly rapid response to injected nuclei and to overexpressed factors. New gene expression can be seen within one or two days at 18°C and this is equivalent to 10 hours at 37°C for mammalian cultured cells. An added advantage of oocytes is that DNA replication is not involved in any gene response observed.

There were three cell lines used for Oocyte-NT and they represent different cell types, namely B10 CHD4 for mESCs (mouse embryonic stem cells), sixiFM for MEFs (mouse embryonic fibroblasts) and C2C12 for mMyos (mouse adult myoblasts) (Figure 5.1.1). These cells were SLO-permeabilized and were injected into GV of *Xenopus* oocytes at Day 0. Two hours later, BrUTP was injected into oocytes, which was incorporated into newly synthesized transcripts after RNA Polymerase II on chromatin of donor cells was replaced by RNA Polymerase II of *Xenopus* oocytes¹³⁶. The oocyte

samples were then collected at Day 1 or Day 2 after Oocyte-NT and processed for RNA-seq.

In order to evaluate the genome-wide effect of oocyte factors, Oocyte-NT was performed, followed by BrUTP injection. Reprogrammed transcriptomes of mESCs, MEFs and mMyos (mESC-NT, MEF-NT and mMyo-NT) were analyzed by RNA-seq (Figure 5.1.1). The time-dependent effect of oocyte factors on MEFs was examined by comparing samples collected at Day 1 and Day 2 after Oocyte-NT (Figure 5.1.1.A). Moreover, the difference among mESC-NT, MEF-NT and mMyo-NT were evaluated by comparing Oocyte-NT samples injected with different cell types and samples were collected at Day 2 after Oocyte-NT (Figure 5.1.1.B).

The RNA-seq data was first analyzed by Angela Simeone for alignment, hierarchical clustering, MDS (multidimensional scaling) analysis and DE (differential expression) analysis. Then, I further analysed these data for Venn diagram, Pearson correlation coefficient, Gene ontology and KEGG pathway. The description of preliminary data validation is in Appendix IV (page 327).

Figure 5.1.1

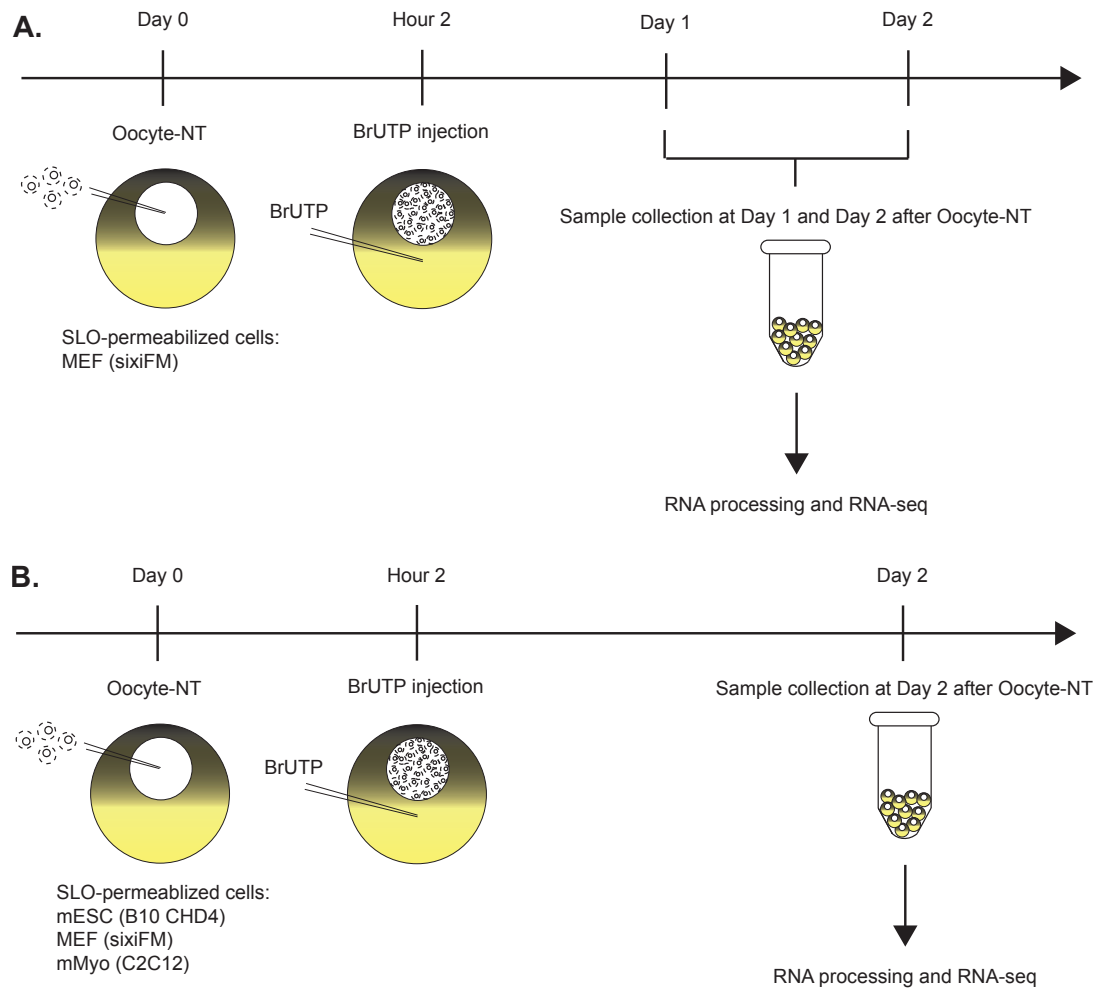


Figure 5.1.1 Sample preparation for evaluating the effect of oocyte factors on various cell types during SCNR by oocytes.

(A) Time-dependent effect of oocyte factors on MEFs at Day 1 and Day 2 after Oocyte-NT

(B) Effects of oocyte factors on mESCs, MEFs and mMyos at Day 2 after Oocyte-NT

5.2 *Xenopus* oocytes reprogram MEFs to a steady state within two days after Oocyte-NT

In chapter 4, it has been shown that oocyte factors up-regulate some pluripotency genes time-dependently by either activating them from a silent state (mOct4, Figure 4.2.1.A, page 110) or by enhancing their expression when they are already expressed in the donor cells during SCNR by oocytes (Figure 4.2.1.B-E, page 110; Figure 4.2.2, page 111). The up-regulation of these oocyte-affected genes in MEFs has been shown to be strong for selected pluripotency genes with a fold change of more than 4 for mSall4, mKlf4, mUtf1 and more than 100 for mSox2 from Day 0 to Day 2 after Oocyte-NT (Figure 4.2.1.B-E, page 110).

Here, in Chapter 5, I compare MEF-NT between Day 1 and Day 2 after Oocyte-NT genome-wide in order to know how many genes are affected by oocyte factors and how strong and effective oocyte factors are in terms of regulating downstream genes.

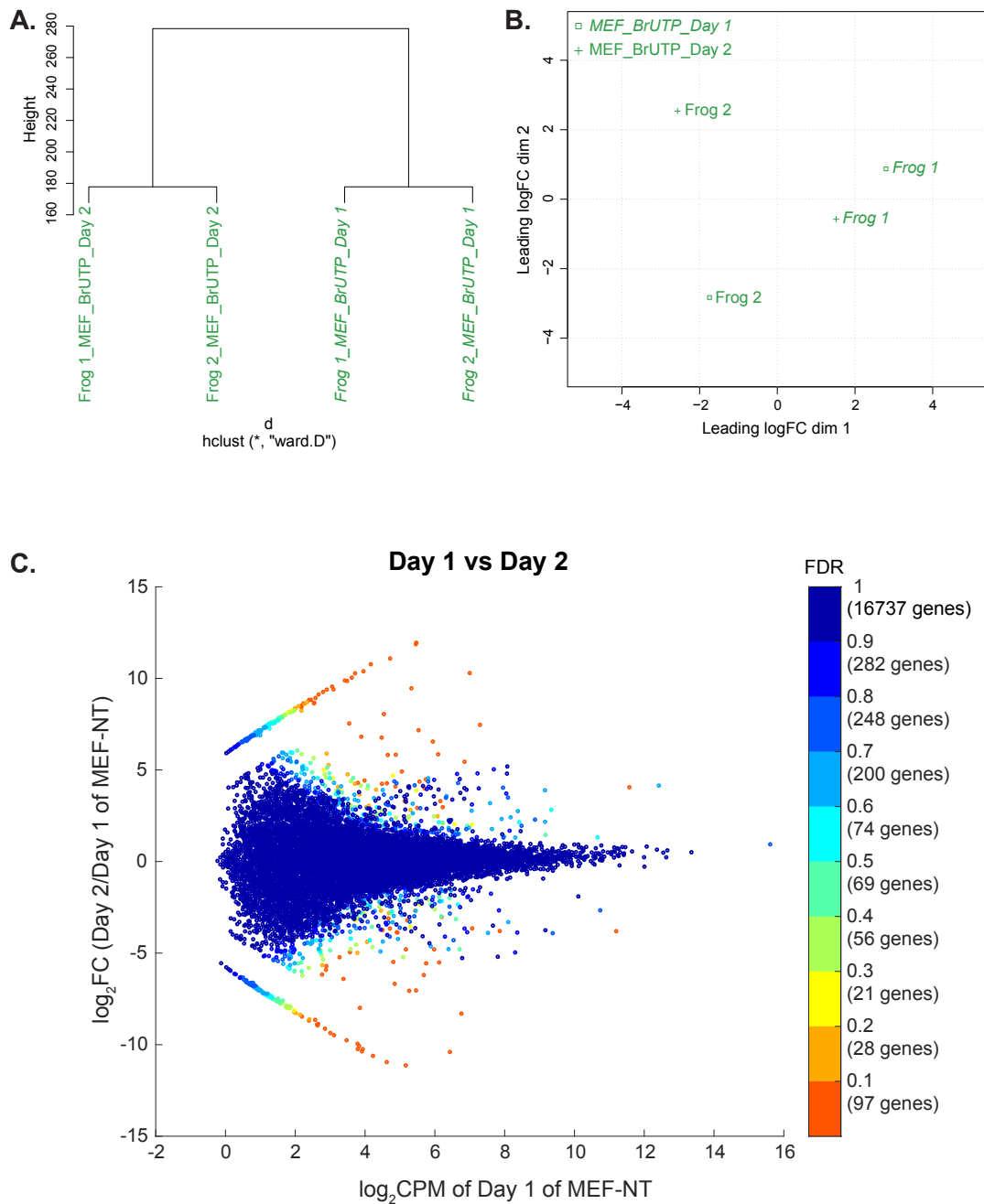
5.2.1 The Day 1 and Day 2 samples of MEF-NT are similar and the expression of only 0.5% of analyzed genes is changed by *Xenopus* oocytes from Day 1 to Day 2 after Oocyte-NT

To examine the difference of libraries between Day 1 and Day 2 samples of MEF-NT, hierarchical clustering, MDS analysis and DE analysis were applied (paired biological duplicates, Frog 1 and 2, Figure 5.2.1). The hierarchical clustering shows that RNA-seq libraries of MEF-NT samples of Frog 1 and 2 collected at Day 2 after Oocyte-NT are grouped together by the expression of

genes in MEF-NT and separated from libraries of MEF-NT samples collected at Day 1 (Figure 5.2.1.A). However, when the data were analyzed via MDS analysis, the RNA-seq libraries of Day 1 samples are not separated from libraries of Day 2 samples and it indicates the difference between Day 1 and Day 2 of MEF-NT is not significant (Figure 5.2.1.B).

When comparing the expression of each gene between the MEF samples collected at Day 1 and Day 2 after Oocyte-NT via DE analysis, there are 97 Oocyte-DE genes regulated by oocyte factors ($FDR < 0.1$, Figure 5.2.1.C). Among these Oocyte-DE genes, 51 Oocyte-DE genes are up-regulated by a fold change of more than 4 and 46 Oocyte-DE genes are down-regulated by the fold change of less than -8 by oocyte factors from Day 1 to Day 2 after Oocyte-NT (Figure 5.2.1.C). Since 17812 newly synthesized transcripts are analyzed by DE analysis and there are 97 Oocyte-DE genes, only 0.5% of newly synthesized transcripts are significantly regulated by oocyte factors from Day 1 to Day 2 after Oocyte-NT. Therefore, the difference of MEF-NT between Day 1 and Day 2 samples is small and gene expression in MEF-NT is not greatly changed by oocyte factors from Day 1 to Day 2 after Oocyte-NT.

Figure 5.2.1



The expression of 51 Oocyte-DE genes in MEF-NT is up-regulated from Day 1 to Day 2 after Oocyte-NT ($FDR < 0.1$, $2 < \log_2FC < 12$)

The expression of 46 Oocyte-DE genes in MEF-NT is down-regulated from Day 1 to Day 2 after Oocyte-NT ($FDR < 0.1$, $-12 < \log_2FC < -3$)

Figure 5.2.1 The difference between MEF-NT at Day 1 and Day 2 after Oocyte-NT is small and only 0.5% analyzed genes are Oocyte-DE genes ($FDR < 0.1$, $n=2$, paired).

(A-B) Hierarchical clustering (A) and MDS analysis (B) show the difference between MEF-NT at Day 1 (in green and *italic*) and MEF-NT at Day 2 (in green) after Oocyte-NT. Height represents the dissimilarity across samples.

(C) DE analysis shows 97 Oocyte-DE genes are significantly up- or down-regulated by oocyte factors by more than 4-fold and 8-fold from Day 1 to Day 2 after Oocyte-NT (orange dots).

5.2.2 The expression of 66% newly synthesized transcripts in MEF-NT remains constant from Day 1 to Day 2 after Oocyte-NT and some pluripotency genes are up-regulated by oocyte factors within one day after Oocyte-NT

To further analyze the data of DE analysis without considering the statistical significance, the \log_2FC between Day 1 and Day 2 samples of MEF-NT against the number of genes is shown as a bar chart and the values for \log_2FC and the number of genes are shown (Figure 5.2.2 and Table 5.2). There are 11631 genes newly synthesized after Oocyte-NT in DE analysis are not changed from Day 1 to Day 2 after Oocyte-NT and it accounts for 66% of analyzed genes ($-1 < \log_2FC < 1$, Figure 5.2.2 and Table 5.2). Additionally, 19% of analyzed genes are weakly up- or down-regulated by oocyte factors ($1 < \log_2FC < 2$ and $-2 < \log_2FC < -1$) and 15% of analyzed genes are strongly up- or down-regulated by oocyte factors ($\log_2FC > 2$ and $\log_2FC < -2$) between Day 1 and Day 2 after Oocyte-NT (Figure 5.2.2 and Table 5.2).

In Chapter 4, some pluripotency genes in MEFs have been shown to be up-regulated by oocyte factors between Day 0 and Day 2 after Oocyte-NT (Figure 4.2.1 and 4.2.2, Page 110 and 111). Now I took them as examples for the comparison between results obtained from QPCR and RNA-seq plus BrUTP. By comparison, some of the pluripotency genes up-regulated by oocyte factors from Day 0 and Day 2 in Chapter 4 (Figure 4.2.1 and 4.2.2, Page 110 and 111) are also up-regulated from Day 1 to Day 2 in this chapter but some are not (Table 5.2). Notably, none of them are judged as Oocyte-DE genes regulated by oocyte factors from Day 1 to Day 2 because all of them have $FDR > 0.1$.

Among these pluripotency genes, the expression of mUtf1 and mEsrrb are enhanced by oocyte factors by 14-fold and 2.2-fold from Day 0 to Day 2 after Oocyte-NT, analyzed by QPCR in Chapter 4 (Figure 4.2.2, Page 111) and the expression of mUtf1 and mEsrrb increase by ~4-fold and ~128-fold from Day 1 to Day 2 after Oocyte-NT, analyzed by RNA-seq plus BrUTP pulldown in this chapter (Table 5.2).

For mSox2, mKlf4 and mSall4, they has been shown to be up-regulated by oocyte factors by 193-fold, 6.5-fold and 9.1-fold from Day 0 to Day 2 after Oocyte-NT in Chapter 4 (Figure 4.2.1.B-D, Page 110) but they are not up-regulated by oocyte factors from Day 1 to Day 2 after Oocyte-NT in this chapter ($-1 < \log_2 FC < 1$, Table 5.2). Therefore, the expression of mSox2, mKlf4 and mSall4 remains constant from Day 1 to Day 2 after Oocyte-NT and the strong enhancement of expression by oocyte factors happens within one day after Oocyte-NT.

Intriguingly, mOct4 in MEFs has been shown to be up-regulated by oocyte factor by 92-fold from a silent or low expressed state from Day 0 to Day 2 after Oocyte-NT in Chapter 4 (Figure 4.2.1.A and 4.3.A, Page 110 and 118) but the expression of mOct4 seems to be down-regulated by oocyte factors from Day 1 to Day 2 after Oocyte-NT in this chapter (Table 5.2). It is possible that the resistance of mOct4 is so strong that the activation of mOct4 is random by oocyte factors and resulting the down-regulation of mOct4 observed here.

Overall, the up-regulation of different pluripotency genes in MEFs at different time points may relate to when and how oocyte factors affect individual gene due to their different chromatin status in donor cells.

Figure 5.2.2

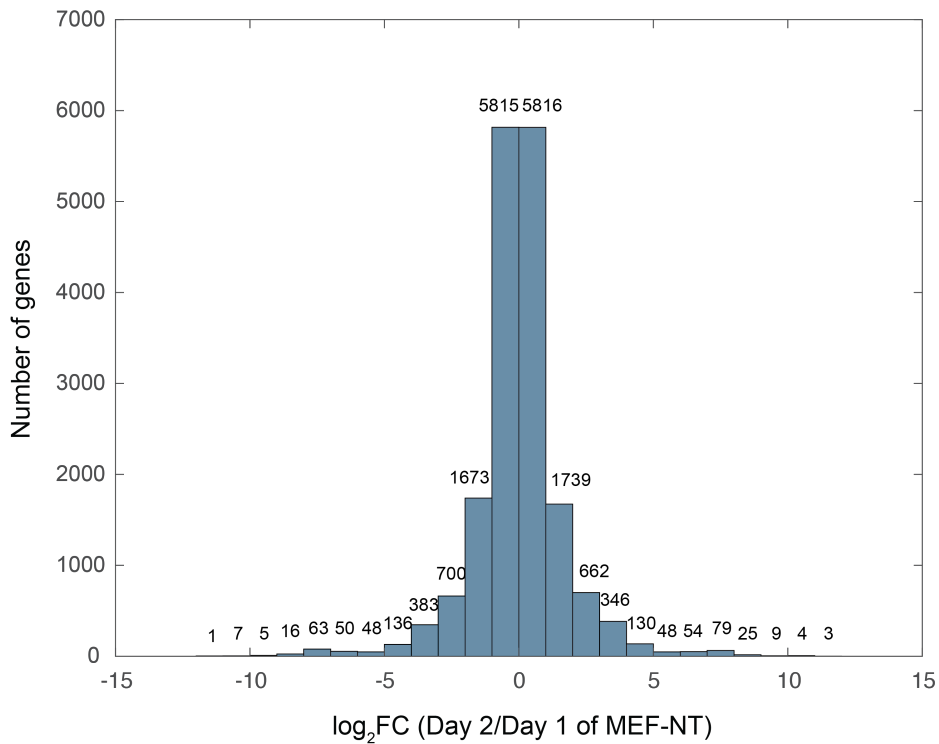


Figure 5.2.2 Without considering statistical significance, 2/3 analyzed genes are not changed by oocyte factors from Day 1 to Day 2 after Oocyte-NT.

The exact number of genes is also shown in Table 5.2. 17812 genes are analyzed. The expression of 66% newly synthesized transcripts is not changed (11631 genes, $-1 < \log_2 \text{FC} < 1$). 10% newly synthesized transcripts are weakly up-regulated (1739 genes, $1 < \log_2 \text{FC} < 2$) and 8% are strongly up-regulated (1360 genes, $\log_2 \text{FC} > 2$) by oocyte factors. 9% newly synthesized transcripts are weakly down-regulated (1673 genes, $-2 < \log_2 \text{FC} < -1$) and 8% are strongly down-regulated (1409 genes, $\log_2 \text{FC} < -2$) by oocyte factors.

Table 5.2

log ₂ FC	Number of genes	% of analysed genes	Pluripotency genes*
11~12	3	0.02%	
10~11	4	0.02%	
9~10	9	0.1%	
8~9	25	0.1%	
7~8	79	0.4%	mEsrrb ⁺
6~7	54	0.3%	
5~6	48	0.3%	
4~5	130	1%	
3~4	346	2%	
2~3	662	4%	mUtf1 ⁺⁺
1~2	1739	10%	
0~1	5816	33%	mKlf4 ⁺⁺ , mSox2 ⁺⁺
-1~0	5815	33%	mKlf2 [^] , mSall4 ⁺⁺
-2~-1	1673	9%	
-3~-2	700	4%	
-4~-3	383	2%	mOct4 ⁺⁺
-5~-4	136	1%	
-6~-5	48	0.3%	
-7~-6	50	0.3%	
-8~-7	63	0.4%	
-9~-8	16	0.1%	
-10~-9	5	0.03%	
-11~-10	7	0.04%	
-12~-11	1	0.01%	

Table 5.2 mEsrrb and mUtf1 are up-regulated by oocyte factors from Day 1 to Day 2 after Oocyte-NT while all the listed pluripotency genes except for mKlf2 are up-regulated by oocyte factors from Day 0 to Day 2 after Oocyte-NT.

* All the listed pluripotency genes in MEF-NT are not significantly regulated by oocyte factors from Day 1 to Day 2 after Oocyte-NT with FDR>0.1, judged by DE analysis.

⁺ mEsrrb in MEFs is up-regulated by oocyte factors by 2.2-fold from Day 0 to Day 2 after Oocyte-NT, judged by QPCR (Figure 4.2.2.B, Page 111).

⁺⁺ mUtf1, mKlf4, mSox2, mSall4 and mOct4 in MEFs are up-regulated by oocyte factors by more than more than 6.5-fold and up to 193-fold from Day 0 to Day 2 after Oocyte-NT, judged by QPCR (Figure 4.2.1 and 4.2.2.A, Page 110 and 111).

[^] As a constant gene, the expression of mKlf2 in MEFs is not changed by oocyte factors from Day 0 to Day 2 after Oocyte-NT.

5.2.3 The batch effects affect the speed of SCNR by *Xenopus* oocytes but different batches of oocytes still can reprogram transcriptomes of MEFs to a highly similar oocyte-steady state at Day 2 after Oocyte-NT

To investigate the oocyte effects on SCNR at different time points, the expression level of newly synthesized transcripts (FPKM) of MEF-NT collected at different days after Oocyte-NT or from different frogs is analyzed by Pearson correlation coefficient (Figure 5.2.3 and 5.2.4). While the value of Pearson correlation coefficient (ρ) approaches 1, genes of MEF-NT from different samples are linearly correlated with an upward trend.

The result demonstrates that MEF-NT of Frog 1 between Day 1 and Day 2 Oocyte-NT samples ($\rho=0.95$, Figure 5.2.3.A) is more correlated linearly than MEF-NT of Frog 2 between Day 1 and Day 2 Oocyte-NT samples ($\rho=0.72$, Figure 5.2.3.B). Hence, MEF-NT of Frog 1 does not change as much as MEF-NT of Frog 2 and MEF-NT of Frog 1 almost reaches the same steady state as at Day 1 after Oocyte-NT when MEF-NT of Frog 2 still changes from Day 1 to Day 2 after Oocyte-NT. A similar result is also shown via MDS analysis when MEF-NT of Day 1 and Day 2 samples of Frog 1 are closer than MEF-NT of Day 1 and Day 2 samples of Frog 2 (Figure 5.2.1.B).

Additionally, Day 1 samples of MEF-NT of Frog 1 and Frog 2 are compared and Day 2 samples of MEF-NT of Frog 1 and Frog 2 are also compared for evaluating the effects of oocyte factors from different batches of oocytes on SCNR at different sample collection days (Figure 5.2.4). For the correlation of Day 1 and Day 2 samples of MEF-NT between Frog 1 and Frog 2, MEF-NT of

Day 2 samples between Day 1 and Day 2 ($\rho=0.98$, Figure 5.2.4.B) are more linearly correlated than Day 1 samples ($\rho=0.71$, Figure 5.2.4.A). Therefore, the expression of genes of MEF-NT of Day 1 samples is slightly different with a lower value of Pearson correlation coefficient than Day 2 samples due to the batch effects of oocytes from different frogs.

MEF-NT of Frog 1 does not change much from Day 1 to Day 2 ($\rho=0.95$, Figure 5.2.3.A) and MEF-NT of Day 2 samples between Frog 1 and Frog 2 are fairly similar ($\rho=0.98$, Figure 5.2.4.B). Therefore, the change of MEF-NT of Frog 2 from Day 1 to Day 2 ($\rho=0.72$, Figure 5.2.3.B) makes the final MEF-NT of Frog 2 reach the same steady state as MEF-NT of Frog 1, which does not change much from Day 1 to Day 2 ($\rho=0.95$, Figure 5.2.3.A). It suggests that the batch effects of oocytes from different female frogs may result in the difference of MEF-NT at Day 1 because different batches of oocytes reprogram cell nuclei with different speeds but it is sufficient for oocyte to reprogram cell nuclei to a steady state within 2 days regardless of the batch effects.

Figure 5.2.3

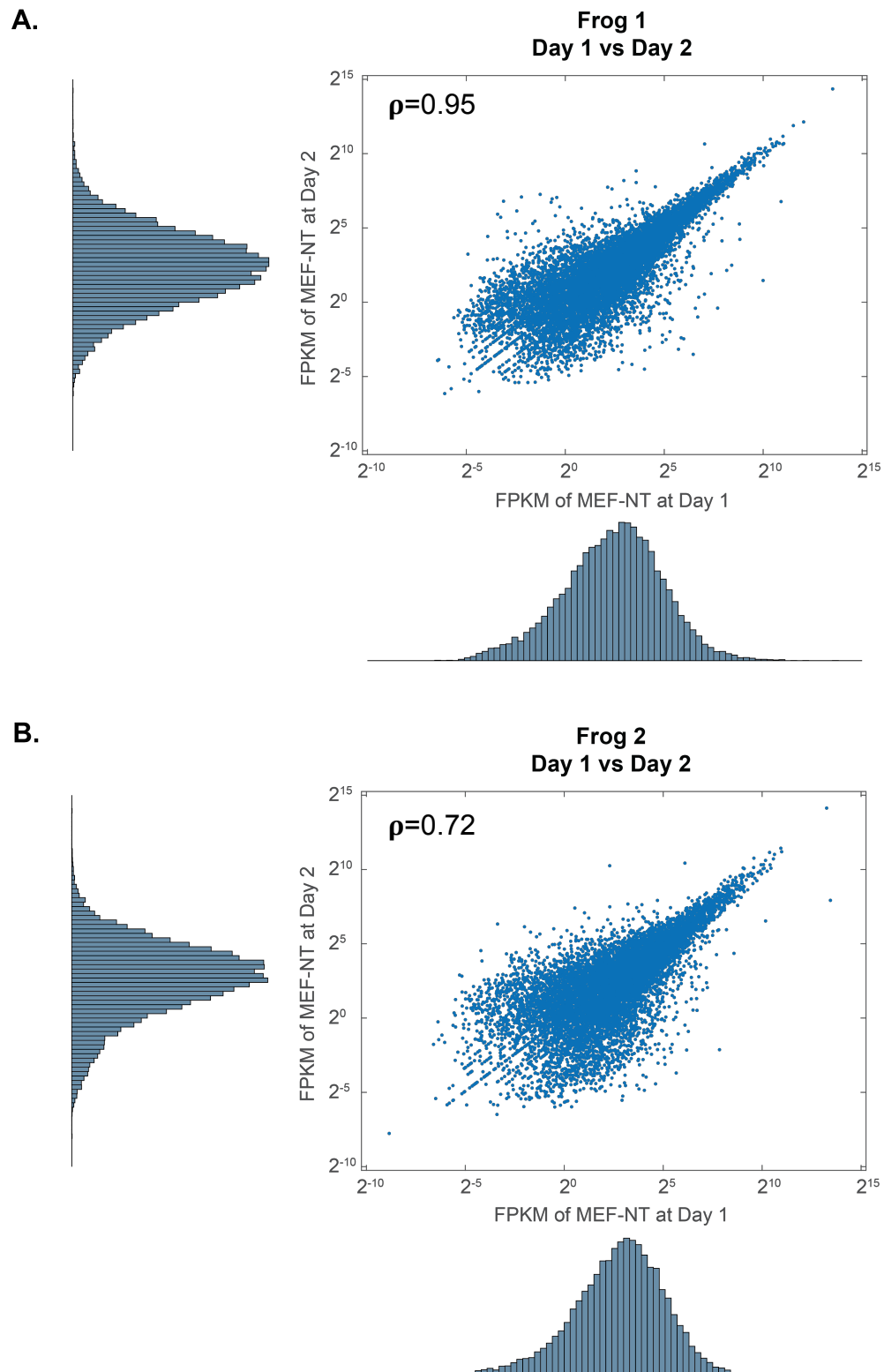


Figure 5.2.3 Pearson correlation coefficient shows different batches of oocytes (A and B) reprogram MEFs to the oocyte-steady state with different speed from Day 1 to Day 2 after Oocyte-NT although the difference caused by batch effects is small.

Figure 5.2.4

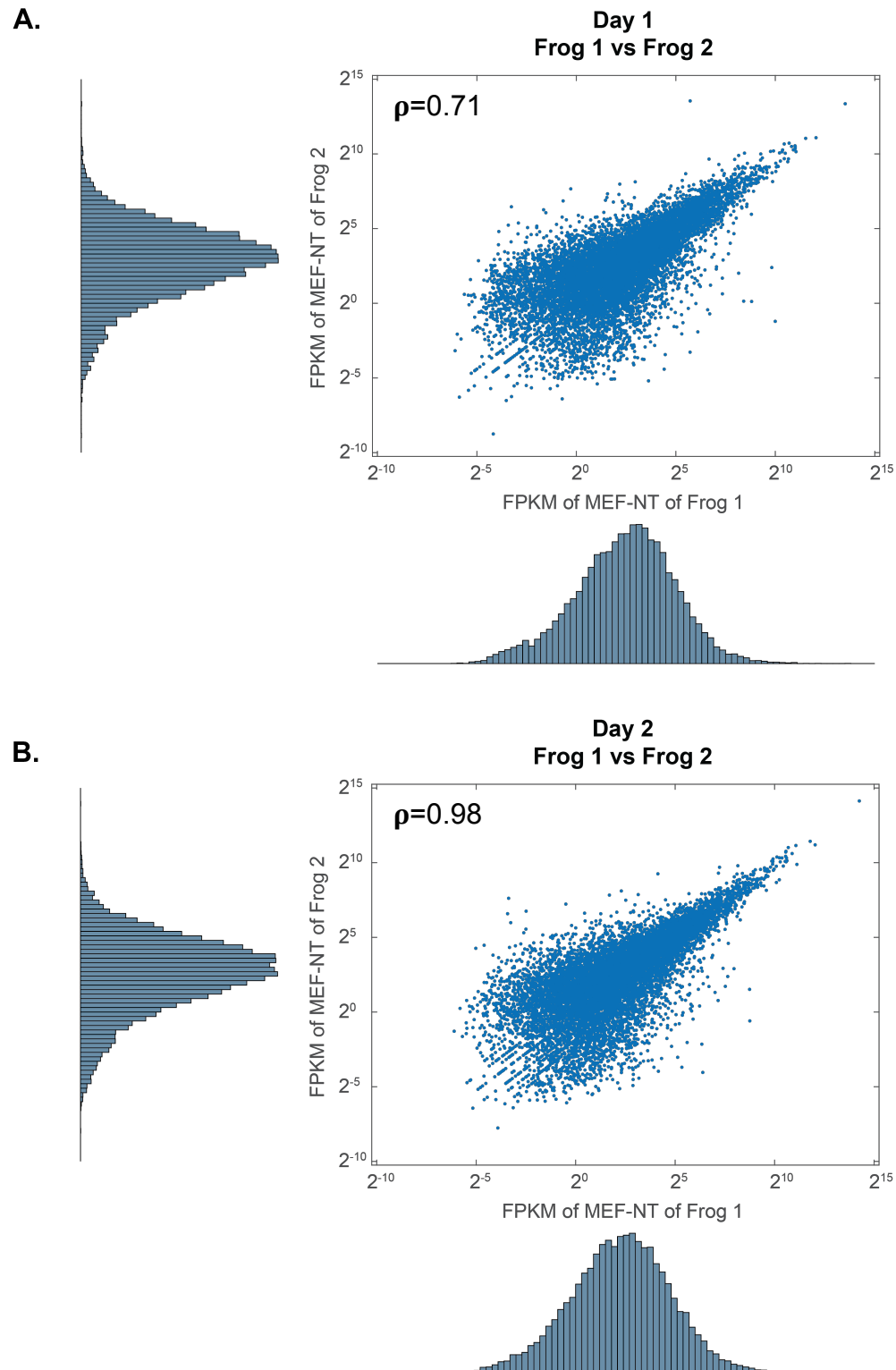


Figure 5.2.4 Pearson correlation coefficient shows batch effects at Day 1 after Oocyte-NT (A) between Frog 1 and Frog 2 is diminished at Day 2 after Oocyte-NT (B) while the expression of newly synthesized transcripts in MEF-NT between Frog 1 and Frog 2 at Day 2 ($\rho=0.98$) is more similar than Day 1 ($\rho=0.71$).

5.2.4 Summary

Since the change made by oocyte factors is time-dependent as has been shown in Chapter 4 (Figure 4.2.1 and 4.2.2, Page 110 and 111), the comparison of samples at different time points could provide clues regarding the way by which oocyte factors of oocytes can successfully reprogram somatic cell nuclei and how resistance of somatic cell nuclei hinders successful reprogramming by oocytes. Therefore, I evaluated the time-dependent difference of MEF-NT by comparing the oocyte samples collected at Day 1 and Day 2 after Oocyte-NT and analyzed them by hierarchical clustering, MDS analysis, DE analysis and Pearson correlation coefficient.

Hierarchical clustering shows that MEF-NT of Day 1 samples is separated from MEF-NT of Day 2 samples (Figure 5.2.1.A). However, MEF-NT of Day 1 samples and Day 2 samples are not clearly separated from each other via MDS analysis (Figure 5.2.1.B). Furthermore, DE analysis shows that only 0.5% of analyzed genes are significantly changed from Day 1 to Day 2 after Oocyte-NT (Figure 5.2.1.C). Hence, the difference between MEF-NT of Day 1 and Day 2 samples is small.

To look further into the data of DE analysis without considering the statistical significance, 66% of analyzed genes are not changed (11631 genes, $-1 < \log_2FC < 1$), 19% of analyzed genes are weakly up- or down-regulated (1739 genes, $1 < \log_2FC < 2$; 1673 genes, $-2 < \log_2FC < -1$) and 15% of analyzed genes are strongly up- or down-regulated (1360 genes, $\log_2FC > 2$; 1409 genes, $\log_2FC < -2$) by oocyte factors from Day 1 to Day 2 after Oocyte-NT (Table 5.2).

Compared to pluripotency genes up-regulated by oocyte factors from Day 0 to Day 2 after Oocyte-NT in Chapter 4 (Figure 4.2.1 and 4.2.2, Page 110 and 111), the expression of some pluripotency genes keeps increasing (mUtf1 and mEsrrb) but some pluripotency genes remains constant (mSox2, mKlf4 and mSall4) from Day 1 to Day 2 after Oocyte-NT in this chapter (Table 5.2). The time period difference indicates that some genes (mSox2, mKlf4 and mSall4) are up-regulated by oocyte factors to a steady state within one day after Oocyte-NT and some genes (mUtf1 and mEsrrb) are up-regulated by oocyte factors later. This is probably due to different epigenetic barriers for each gene and the stepwise molecular mechanisms to remove these barriers after nuclear transfer⁵.

When comparing MEF-NT of Day 1 and Day 2 samples reprogrammed by two batches of *Xenopus* oocytes, Pearson correlation coefficient shows that one batch of oocytes can reprogram MEFs to a steady state less than a day without much change of transcriptomes from Day 1 to Day 2 (Figure 5.2.3.A). The other batch of oocytes needs more than a day to reprogram MEFs to a steady state (Figure 5.2.3.B and 5.2.4.B). The batch effects of oocytes decrease while increasing the time for oocyte reprogramming MEFs from one day to two days (Figure 5.2.4). Thus, the batch effects slightly affect the speed of transcriptional reprogramming by oocytes but the transcriptomes of reprogrammed cell nuclei would end up at a steady state built up by oocyte factors.

All in all, the majority of the newly synthesized transcripts are reprogrammed successfully by oocyte factors to a certain level within a day and their expression remains constant from Day 1 to Day 2 after Oocyte-NT. Some genes are still changed from Day 1 to Day 2 after Oocyte-NT caused by the batch effects and the resistance of these genes. When increasing the reprogramming time from one day to two day, the batch effects can be diminished since batch effects affect the speed of SCNR by oocytes. For MEFs, two days are sufficient for *Xenopus* oocytes to reprogram the transcriptomes of MEFs to a steady state.

5.3 *Xenopus* oocytes reprogram the transcriptomes of mESCs, MEFs and mMyos to be almost identical at Day 2 after Oocyte-NT

During SCNR by eggs, the efficiency of successful reprogramming is usually low due to the barriers of transplanted cell nuclei². In 5.2, it has been shown that transcriptomes of MEFs can be reprogrammed by oocyte factors to a steady state within two days regardless of batch effects of *Xenopus* oocytes. It seems MEFs can be reprogrammed successfully by oocyte factors within two days after Oocyte-NT but how about other cell types? Do the original transcriptional machineries of donor cell nuclei affect the SCNR by oocytes? Do the chromatin structures of donor cell nuclei affect SCNR by oocytes?

Here, I used three cell types for Oocyte-NT and evaluated the difference between reprogrammed transcriptomes of mESCs (mESC-NT), MEFs (MEF-NT) and mMyos (mMyo-NT) and the difference may relate to their potential functions. Since these experiments were designed to be paired for examining the effect of *xklf2*-HA on SCNR by oocytes in Chapter 6, two in three of the comparisons of reprogrammed transcriptomes via DE analysis are not paired in this section (mMyo-NT versus mESC-NT, paired, Figure 5.3.1; MEF-NT versus mMyo-NT, MEF-NT versus mESC-NT, unpaired, Figure 5.3.2 and 5.3.3). However, the results of paired and unpaired analyses are similar and the batch effects for SCNR by oocytes at Day 2 after Oocyte-NT may not be significant, as has been shown in section 5.2.

5.3.1 mESC-NT, MEF-NT and mMyo-NT are not grouped by the difference of cell types or different origins of oocyte batches

The difference between mESC-NT, MEF-NT and mMyo-NT at Day 2 after Oocyte-NT were first evaluated by hierarchical clustering and MDS analysis (Figure 5.3.1.A and B, Figure 5.3.2.A and B and Figure 5.3.3.A and B). The comparison of mMyo-NT to mESC-NT is paired (n=3). The comparisons between MEF-NT and mMyo-NT and MEF-NT and mESC-NT are unpaired (n=3).

For the comparison between mMyo-NT and mESC-NT (paired, n=3), three mMyo-NTs are not grouped together against three mESC-NTs and the batches of oocytes do not affect the grouping (Figure 5.3.1.A and B). For three libraries to the right of hierarchical clustering, mMyo-NT of Frog 3 was first grouped with mESC-NT of Frog 2 and then grouped with mESC-NT of Frog 3 (Right, Figure 5.3.1.A). For the libraries of the left hand branches, mMyo-NTs of Frog 2 and Frog 4 are grouped together first and then grouped with mESC-NT of Frog 4 (Left, Figure 5.3.1.A). Additionally, this shows that mMyo-NTs of Frog 2, 3 and 4 are not separated from mESC-NTs of Frog 2, 3, and 4 via MDS analysis (Figure 5.3.1.B).

Regarding the unpaired comparisons, namely MEF-NTs versus mMyo-NTs and MEF-NTs versus mESC-NTs, similar results were shown that reprogrammed transcriptomes of mESCs, MEFs and mMyos are not grouped or separated by the cell types via hierarchical clustering and MDS analysis (Figure 5.3.2.B and 5.3.3.B).

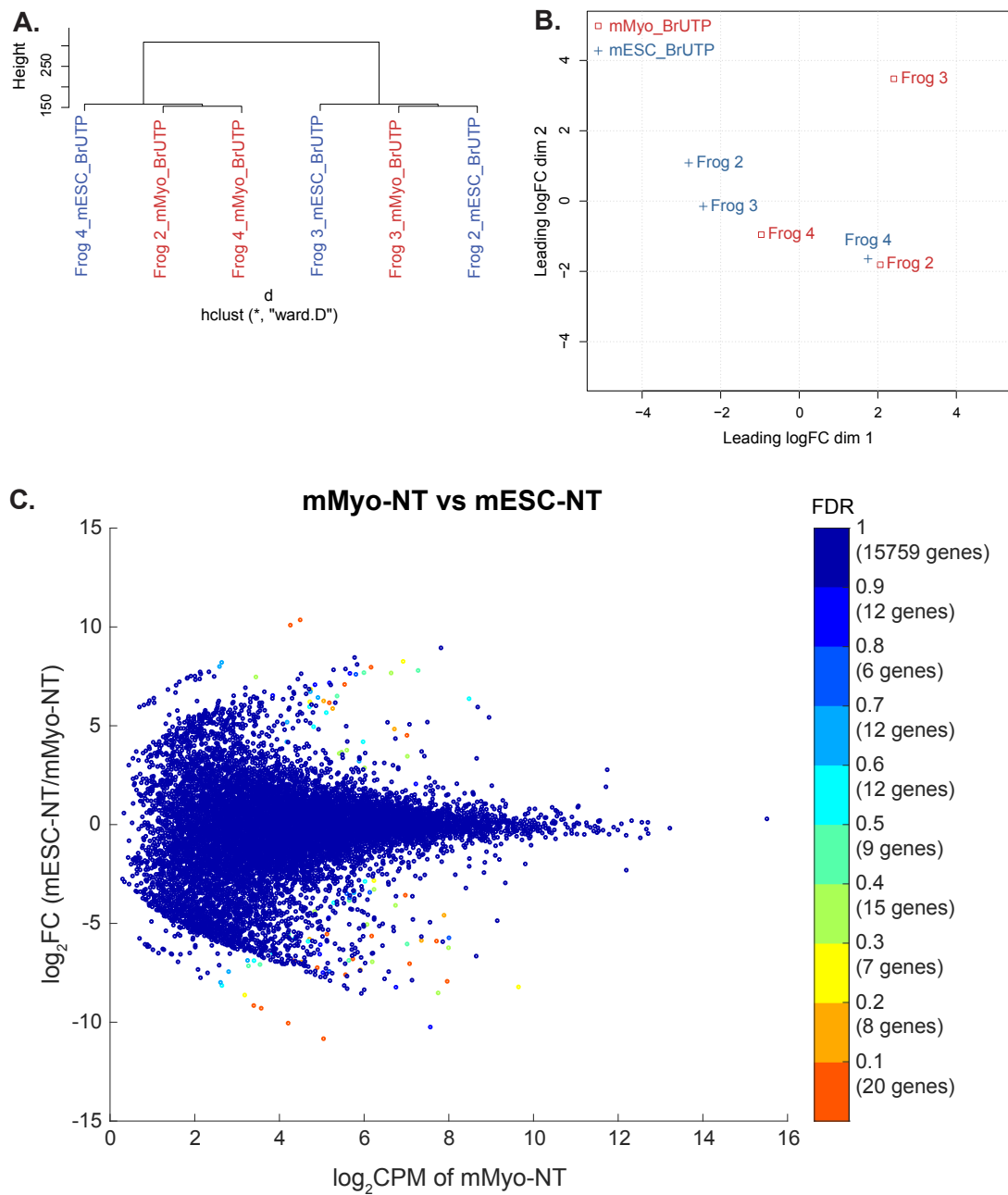
5.3.2 There are less than 40 Cell-type DE genes differentially expressed between reprogrammed transcriptomes of mESCs, MEFs and mMyos at Day 2 after Oocyte-NT

For the DE analysis of mMyo-NT versus mESC-NT, there are 20 Cell-type-DE genes expressed to different levels in mMyo-NT and mESC-NT at Day 2 after Oocyte-NT and this makes up only 0.1% of total analysed genes by DE analysis (FDR<0.1, paired, n=3, Figure 5.3.1.C). Six Cell-type-DE genes between mMyo-NT and mESC-NT are expressed more in mESC-NT than mMyo-NT by more than 16-fold and 14 Cell-type-DE genes are expressed more in mMyo-NT than mESC-NT by more than 8-fold (Figure 5.3.1.C).

Furthermore, for the unpaired DE analyses between MEF-NT and mMyo-NT and MEF-NT and mESC-NT, these demonstrate that less than 40 Cell-type-DE genes are differentially expressed between either MEF-NT and mMyo-NT or MEF-NT and mESC-NT at Day 2 after Oocyte-NT (Figure 5.3.2.C and 5.3.3.C). While there are 34 Cell-type-DE genes that differ between MEF-NT and mMyo-NT, 11 Cell-type-DE genes are more expressed in mMyo-NT than in MEF-NT by more than 16-fold and 23 Cell-type-DE genes are more expressed in MEF-NT than in mMyo-NT by more than 8-fold (Figure 5.3.2.C). On the other hand, there are 24 Cell-type-DE genes differentially expressed between MEF-NT and mESC-NT and 11 Cell-type-DE genes are expressed more in mESC-NT by more than 16-fold and 13 Cell-type-DE genes are expressed more in MEF-NT than the other cell types by more than 8-fold (Figure 5.3.3.C).

Overall, DE analysis shows that less than 0.3% of analyzed genes are Cell-type-DE genes when comparing reprogramming transcriptomes between cell types. In conclusion, DE analysis between cell types indicates that the reprogrammed transcriptomes at Day 2 after Oocyte-NT are almost identical statistically irrespective of the different potencies and functions of donor cells before Oocyte-NT.

Figure 5.3.1



The expression of 6 Cell-type-DE genes is more in mESC-NT than in mMyo-NT (FDR<0.1, $4 < \log_2FC < 11$)

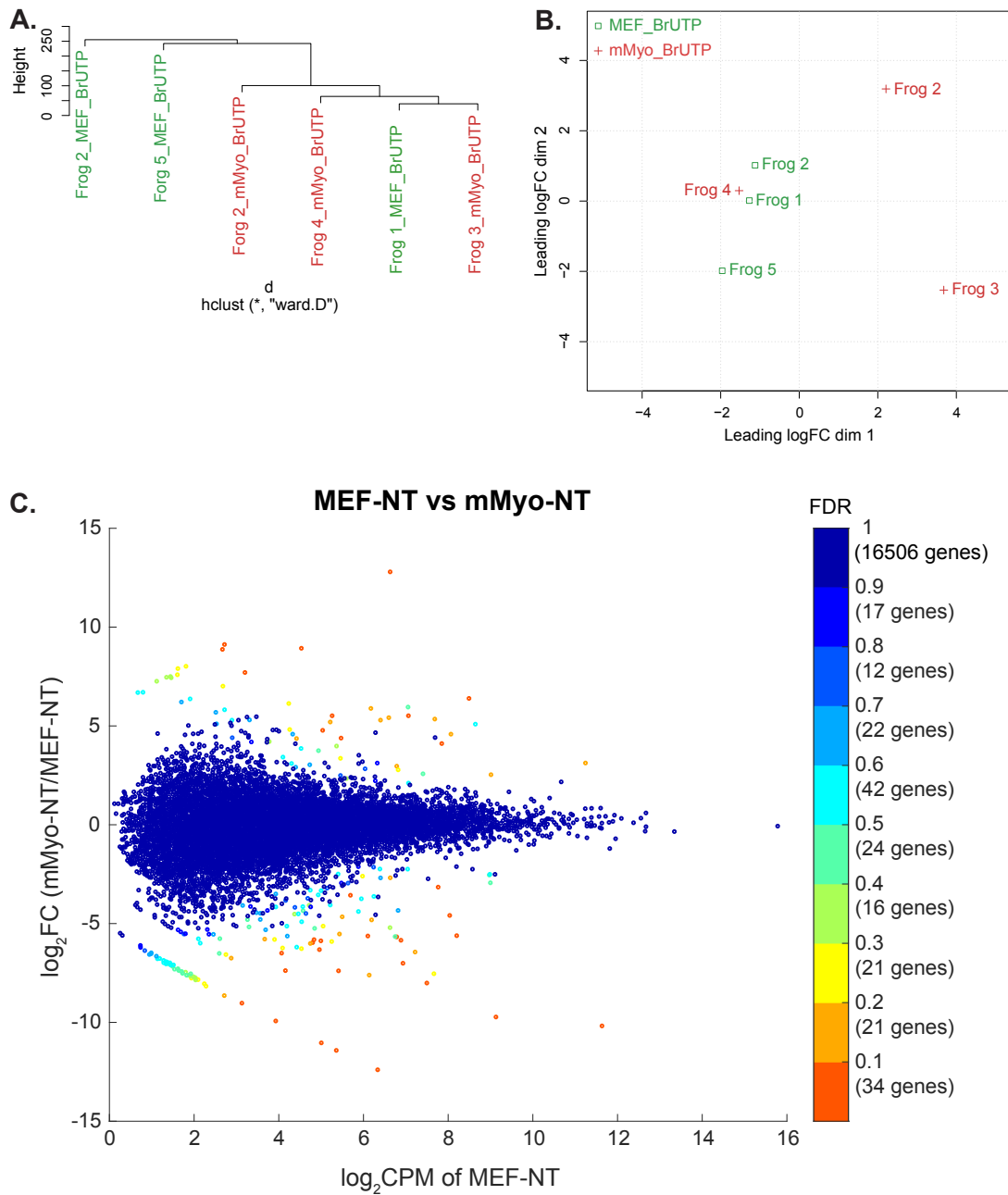
The expression of 14 Cell-type-DE genes is more in mMyo-NT than in mESC-NT (FDR<0.1, $-11 < \log_2FC < -3$)

Figure 5.3.1 The difference between mMyo-NT and mESC-NT is small and only 0.1% of analyzed genes are Cell-type-DE genes (FDR<0.1, n=3, paired).

(A-B) Hierarchical clustering (A) and MDS analysis (B) show the difference between mMyo-NT (in red) and mESC-NT (in blue) at Day 2 after Oocyte-NT. Height represents the dissimilarity across samples.

(C) DE analysis shows 20 Cell-type-DE genes are differentially expressed between mMyo-NT and mESC-NT after Oocyte-NT (orange dots).

Figure 5.3.2



The expression of 11 Cell-type-DE genes is more in mMyo-NT than MEF-NT (FDR<0.1, $4 < \log_2$ FC<13)

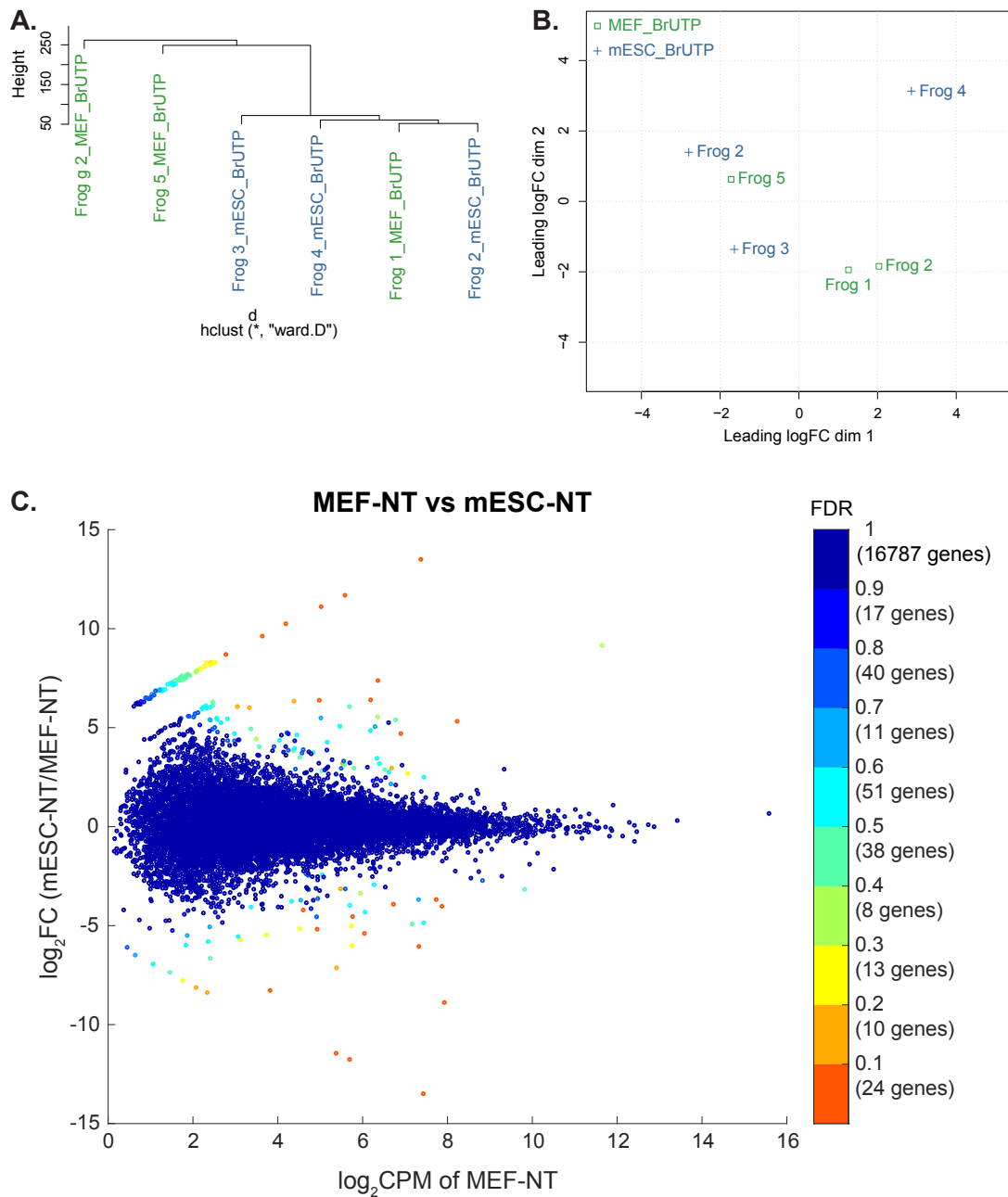
The expression of 23 Cell-type-DE genes is more in MEF-NT than mMyo-NT (FDR<0.1, $-13 < \log_2$ FC<-3)

Figure 5.3.2 The difference between MEF-NT and mMyo-NT is small and only 0.2% of analyzed genes are Cell-type-DE genes (FDR<0.1, n=3, unpaired).

(A-B) Hierarchical clustering (A) and MDS analysis (B) show the difference between MEF-NT (in green) and mMyo-NT (in red) at Day 2 after Oocyte-NT. Height represents the dissimilarity across samples.

(C) DE analysis shows 34 Cell-type-DE genes are differentially expressed between MEF-NT and mMyo-NT after Oocyte-NT (orange dots).

Figure 5.3.3



The expression of 11 Cell-type-DE genes is more in mESC-NT than in MEF-NT (FDR<0.1, $4 < \log_2FC < 14$)

The expression of 13 Cell-type-DE genes is more in MEF-NT than in mESC-NT (FDR<0.1, $-14 < \log_2FC < -3$)

Figure 5.3.3 The difference between MEF-NT and mESC-NT is small and only 0.1% of analyzed genes are Cell-type-DE genes (FDR<0.1, n=3, unpaired).

(A-B) Hierarchical clustering (A) and MDS analysis (B) show the difference between MEF-NT (in green) and mESC-NT (in blue) at Day 2 after Oocyte-NT. Height represents the dissimilarity across samples.

(C) DE analysis shows 24 Cell-type-DE genes are differentially expressed between MEF-NT and mESC-NT after Oocyte-NT (orange dots).

5.3.3 At Day 2 after Oocyte-NT, more than 12000 genes in mESC-NT, MEF-NT and mMyo-NT are reprogrammed by oocyte factors to the same level and, in addition to those genes, 1000-6000 genes are cell-type specific in the donor cells

Less than 0.3% of analysed genes are Cell-type-DE genes when comparing mMyo-NT to mESC-NT, MEF-NT to mMyo-NT, and MEF-NT to mESC-NT. Since more than 99% of analysed genes are non-DE genes, judged by DE analysis, it seems the majority of newly synthesized transcripts of mESC-NT, MEF-NT and mMyo-NT are reprogrammed successfully to the same level by oocyte factors.

However, are these Cell-type-DE genes the only cell-type specific genes? Are there other genes that are cell-type specific in the donor cells and is the expression of these genes after nuclear transfer maintained by transcriptional machineries of donor cells? Are there other genes that are donor cell-type specific and resistant to SCNR by oocytes due to the chromatin structure of these genes in the donor cells?

To investigate these questions, I analysed the same RNA-seq data of mESC-NT, MEF-NT and mMyo-NT by comparing the relative expression of newly synthesized transcripts (FPKM) via Venn diagrams and Pearson correlation coefficients (Figure 5.3.4 and 5.3.5). Reprogrammed transcriptomes for comparison are all transplanted into the same batch of oocytes (Frog 2) so the batch effect can be minimized and the difference between reprogrammed transcriptomes is caused by the cell-type specific traits of donor cells.

When comparing the genes, whose transcripts are newly synthesized after Oocyte-NT or maintained by the transcriptional machineries of donor cells, there are 12070 genes expressed in both mMyo-NT and mESC-NT (FPKM>0, Figure 5.3.4.A). In addition to these genes, 2902 cell-type specific genes are expressed only in mMyo-NT and 3683 cell-type specific genes are expressed only in mESC-NT (FPKM>0, Figure 5.3.4.A). Then, the expression level of those genes expressed in both mMyo-NT and mESC-NT is further analyzed by Pearson correlation coefficient and it shows that the expression level of those genes are linearly correlated ($\rho=0.80$, Figure 5.3.5.A). Since the value of Pearson correlation coefficient (ρ) is close to 1, 12070 genes in both mMyo-NT and mESC-NT are reprogrammed by oocyte factors to highly similar level at Day 2 after Oocyte-NT.

For the comparison of newly synthesized transcripts of MEF-NT to mMyo-NT, 13303 genes are expressed in both MEF-NT and mMyo-NT (Figure 5.3.4.B). Furthermore, 5107 cell-type specific genes are only expressed in MEF-NT and 1669 cell-type specific genes are expressed only in mMyo-NT (Figure 5.3.4.B). For the comparison of MEF-NT to mESC-NT, 13537 genes are expressed in both reprogrammed transcriptomes. 4873 cell-type specific genes are expressed only in MEF-NT and 2216 cell-type specific genes are only expressed in mESC-NT (Figure 5.3.4.C). When comparing the expression level of shared genes between two reprogrammed transcriptomes, the expression level of shared genes between MEF-NT and mMyo-NT is linearly correlated with $\rho=0.81$ and the expression level of shared genes

between MEF-NT and mESC-NT is linearly correlated with $\rho=0.94$ (Figure 5.3.5.B and 5.3.5.C). Therefore, more than 13000 genes are reprogrammed by oocyte factors to highly similar level at Day 2 after Oocyte-NT when comparing MEF-NT with either mMyo-NT or mESC-NT.

Notably, because the value of Pearson correlation coefficient for MEF-NT versus mESC-NT ($\rho=0.94$) is higher than the value for mMyo-NT versus mESC-NT ($\rho=0.80$) and MEF-NT versus mMyo-NT ($\rho=0.81$), the gene expression level of MEF-NT and mESC-NT are reprogrammed by oocyte factors to a more identical level than other two comparisons (Figure 5.3.5). This suggests that embryonic cell types, namely mESCs and MEFs, at early developmental stages may be more susceptible to the effect of oocyte factors than differentiated adult cells (mMyos) and are reprogrammed to the steady state more easily than adult cells when analysed at Day 2 after Oocyte-NT.

Figure 5.3.4

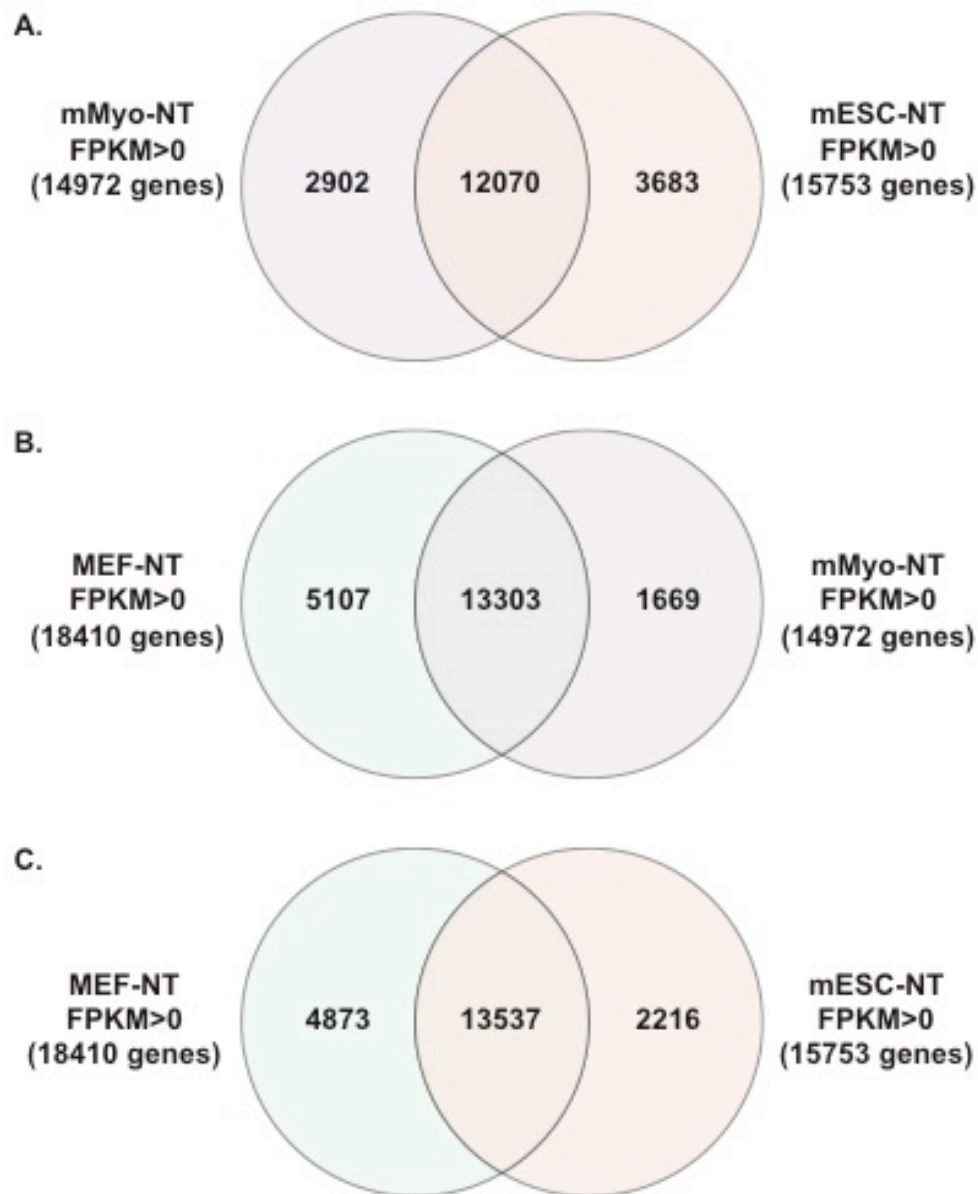


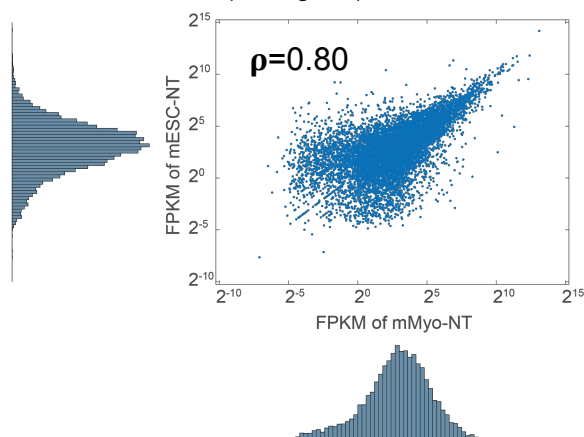
Figure 5.3.4 More than 12000 genes have transcripts synthesized after Oocyte-NT in reprogrammed transcriptomes of all three cell types and 1000-6000 cell-type specific genes have transcripts synthesized after Oocyte-NT in at least one of the three cell types.

Genes with FPKM>0 are those genes have transcripts synthesized after Oocyte-NT.

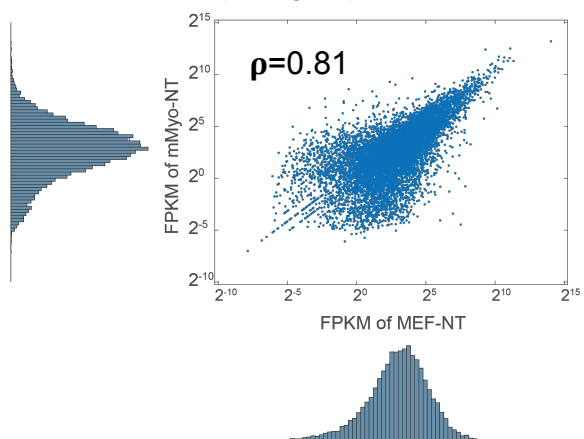
Reprogrammed transcriptomes compared here are all from Frog 2.

Figure 5.3.5

**A. Genes expressed in mMyo-NT vs mESC-NT
(12070 genes)**



**B. Genes expressed in MEF-NT vs mMyo-NT
(13303 genes)**



**C. Genes expressed in MEF-NT vs mESC-NT
(13537 genes)**

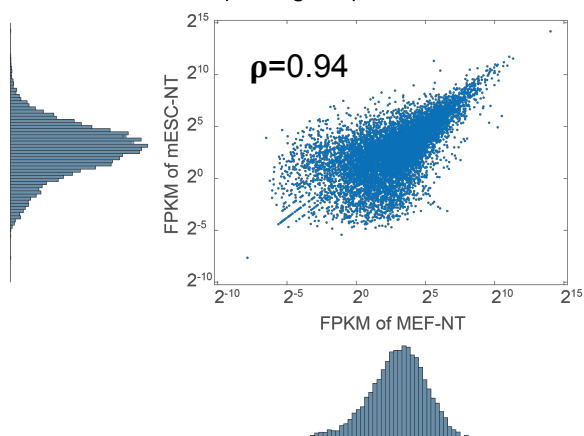


Figure 5.3.5 More than 12000 genes are reprogrammed to a fairly similar level in reprogrammed transcriptomes of all three cell types while expression of genes is highly linearly correlated.

5.3.4 Summary

During SCNR by eggs, the resistance of genes in somatic cell nuclei is the major factor causing the low efficiency of successful reprogramming. In 5.3, it has been shown that *Xenopus* oocytes can reprogram the majority of genes (99%) in MEFs to a steady state within 2 days after Oocyte-NT but it is unknown if the resistant genes are also reprogrammed successfully. In this section, I tried to identify cell-type specific genes by comparing mESC-NT, MEF-NT and mMyo-NT at Day 2 after Oocyte-NT.

RNA-seq libraries of mESC-NT, MEF-NT and mMyo-NT were first analyzed by hierarchical clustering and MDS analysis. These analyses show that libraries of mESC-NT, MEF-NT and mMyo-NT are not grouped by cell types or the batches of oocytes (Figure 5.3.1.A-B, 5.3.2.A-B and 5.3.3.A-B). When comparing mESC-NT, MEF-NT and mMyo-NT via DE analysis, there are less than 0.3% of analyzed genes differentially expressed between different cell types and more than 99% of analyzed genes not differentially expressed (Figure 5.3.1.C, 5.3.2.C, 5.3.3.C).

To look into the difference among mESC-NT, MEF-NT and mMyo-NT, newly synthesized transcripts of reprogrammed transcriptomes of different cell types are evaluated by Venn diagrams and Pearson correlation coefficients. This shows that 1000 to 6000 cell-type specific genes are expressed in mESC-NT, MEF-NT and mMyo-NT and more than 12000 genes in mESC-NT, MEF-NT and mMyo-NT are reprogrammed to the same level by oocyte factors (Figure 5.3.4 and 5.3.5).

The different outcomes for Cell-type-DE genes via DE analysis and cell-type specific genes via Venn diagrams suggests that the definition for those genes in donor cells, which affect SCNR by oocytes, needs to be more delicately interpreted and these cell-type specific genes are discussed further in the following sections.

5.4 Transcriptional reprogramming of mESCs, MEFs and mMyos by *Xenopus* oocytes

In 5.3, it has been shown that *Xenopus* oocytes can reprogram transcriptomes of mESCs, MEFs and mMyos to be very similar in terms of the expression level of newly synthesized transcripts and the majority of these genes are regulated to a steady state by oocyte factors at Day 2 after Oocyte-NT. However, the way by which the original transcriptomes of mESCs, MEFs and mMyos are reprogrammed by oocytes is not known. In this section, I ask how the transcriptomes of mESCs, MEFs and mMyos change before and after Oocyte-NT and what kinds of genes are silenced after Oocyte-NT.

5.4.1 Large-scale change of transcriptomes of mESCs, MEFs and mMyos before and after Oocyte-NT

To examine the change of transcriptomes of mESCs, MEFs and mMyos and understand how oocyte factors regulate genes during SCNR by oocytes, the genes of transcriptomes of mESCs, MEFs and mMyos before and after Oocyte-NT are compared via Venn diagrams and Pearson correlation coefficients (Figure 5.4.1, 5.4.2 and 5.4.3).

The transcriptomes of original mESCs and MEFs refer to database, Expression Atlas (EMBL-EBI). The transcriptome of mMyos refers to database, ENCODE project. Additionally, housekeeping genes are included in these comparisons since they are typically required for maintaining the basic

functions of all cells¹⁴⁴. Including housekeeping genes in the comparisons helps to define the shared and unshared genes before and after Oocyte-NT.

The genes in transcriptomes of mESCs, MEFs and mMyos with FPKM \geq 1 before and after Oocyte-NT are first compared via Venn diagrams (Figure 5.4.1). It shows that 4070 genes in mESCs are silenced and 4718 genes are activated in mESC-NT by oocyte factors (Figure 5.4.1.A). In MEFs, 2591 genes are silenced and 6524 genes are activated by oocyte factors after Oocyte-NT (Figure 5.4.1.B). In mMyos, 4584 genes are silenced and 4508 genes are activated by oocyte factors after Oocyte-NT (Figure 5.4.1.C). For the housekeeping genes, more genes are silenced by oocyte factors than the genes that are activated after Oocyte-NT in all three cell types (Figure 5.4.1).

For the shared genes in transcriptomes with FPKM \geq 1 before and after Oocyte-NT, 8658 genes are shared in mESCs, 8660 genes are shared in MEFs and 7892 genes are shared in mMyos before and after Oocyte-NT (Figure 5.4.1). Among these shared genes, approximate 3000 genes are housekeeping genes (Figure 5.4.1). When examining the expression level of these shared genes, excluding house keeping genes, via Pearson correlation coefficient, the expression level of these genes are lowly linearly correlated with $\rho=0.38$ for mESCs, $\rho=0.07$ for MEFs and $\rho=0.15$ for mMyos between shared genes before and after Oocyte-NT (Figure 5.4.2). Therefore, regardless of housekeeping genes, 5000-6000 shared genes are expressed before and after Oocyte-NT and the expression of these genes is subjected to change to the level determined by oocyte factors.

Furthermore, since highly expressed genes are usually considered to be important to the functions of specified cell types, the top 1000 genes highly expressed in mESCs, MEFs and mMyos before and after Oocyte-NT are compared (Figure 5.4.3). It shows that 771 genes in mESCs are down-regulated by oocyte factors to be less expressed than the level of the top 1000 genes in mESC-NT. Through the same comparison, 831 genes in MEFs and 796 genes in mMyos are down-regulated to be less than top 1000 genes in MEF-NT and mMyo-NT by oocyte factors, respectively (Figure 5.4.3.B and 5.4.3.C). Among these top 1000 genes, 35%~50% are housekeeping genes in transcriptomes of mESCs, MEFs and mMyos both before and after Oocyte-NT (Figure 5.4.3).

Altogether, the large-scale change of transcriptomes contributed by oocyte factors contains both up-regulation and down-regulation of genes. For up-regulation of genes, oocyte factors can either enhance the expression of genes, which are already transcribed in donor cells, or activate genes, which are silent in donor cells. 4718 genes in mESCs, 6524 genes in MEFs and 4508 genes in mMyos are activated by oocyte factors. On the other hand, oocyte factors can also down-regulate genes in donor cells either by repressing the expression of genes to a certain level or by silencing genes if these genes are not needed for reprogrammed transcriptomes. 4070 genes in mESCs, 2591 genes in MEFs and 4584 genes in mMyos are silenced by oocyte factors.

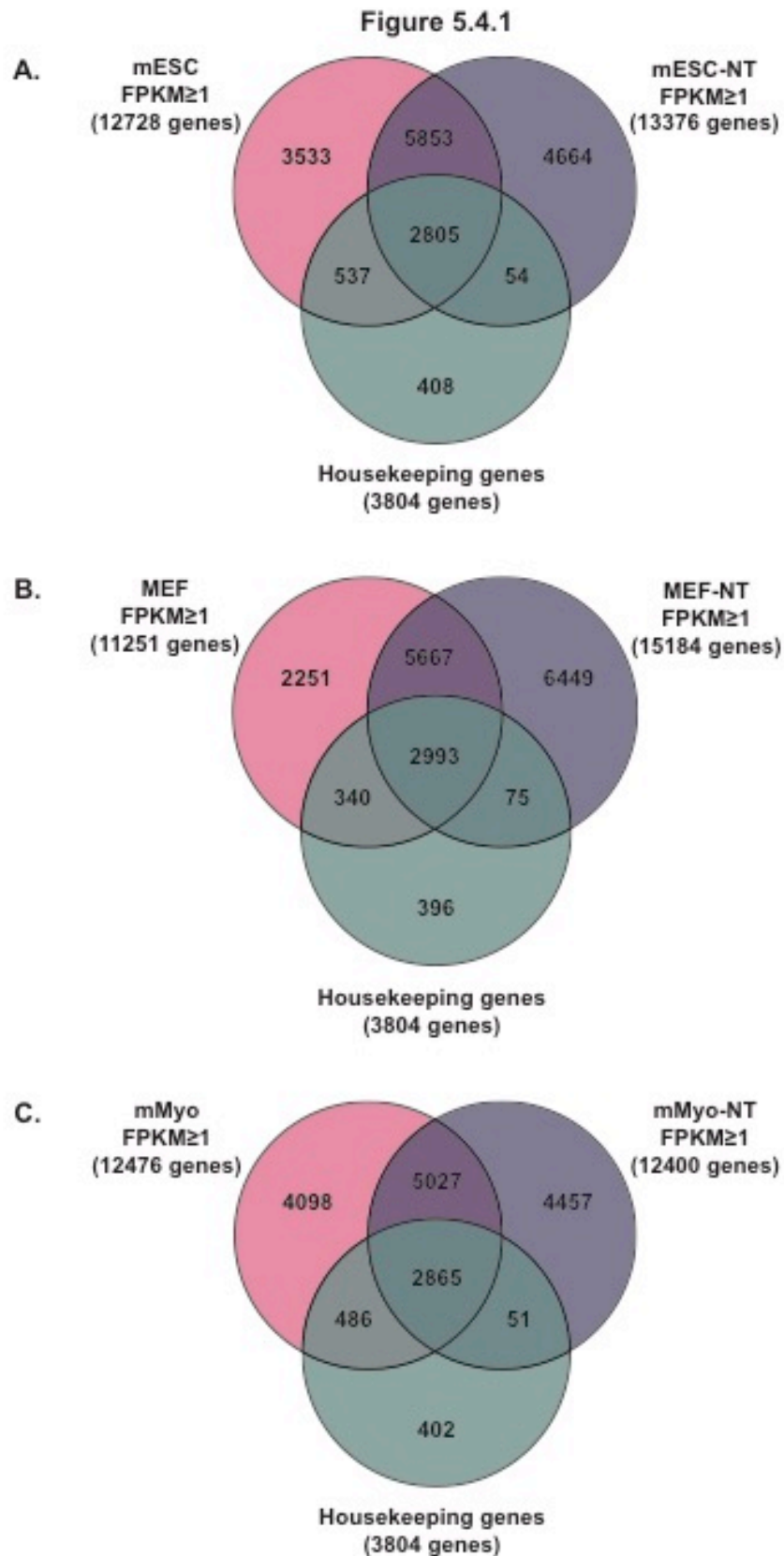


Figure 5.4.1 Large-scale transcriptional change of mESCs, MEFs and mMyos before and after Oocyte-NT.

(Figure legend continues on the next page)

Approximate 1/3 of genes expressed before and after Oocyte-NT are housekeeping genes.

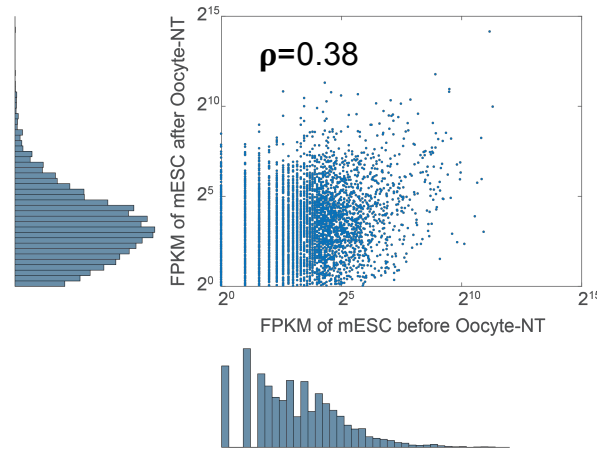
(A.) In mESCs, 8658 genes are expressed both before and after Oocyte-NT and 32% of them are housekeeping genes. Additionally, there are 4070 genes (=3533+537) silenced and 4718 genes (=4664+54) activated after Oocyte-NT. 13% of silenced genes are housekeeping genes and 1% of activated genes are housekeeping genes.

(B.) In MEFs, 8660 genes are expressed both before and after Oocyte-NT and 35% of them are housekeeping genes. Additionally, there are 2591 genes (=2251+340) silenced and 6524 genes (=6449+75) activated after Oocyte-NT. 13% of silenced genes are housekeeping genes and 1% of activated genes are housekeeping genes.

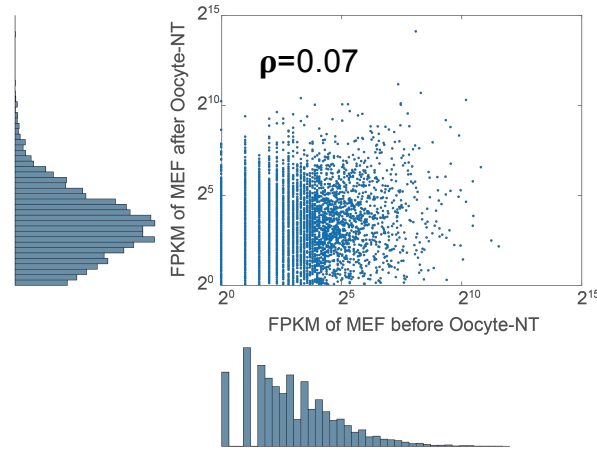
(C) In mMyos, 7892 genes are expressed both before and after Oocyte-NT and 36% of them are housekeeping genes. Additionally, there are 4584 genes (=4098+486) silenced and 4508 genes (=4457+51) activated after Oocyte-NT. 11% of silenced genes are housekeeping genes and 1% of activated genes are housekeeping genes.

Figure 5.4.2

**A. Genes expressed in mESCs before and after Oocyte-NT
(8658 genes)**



**B. Genes expressed in MEFs before and after Oocyte-NT
(8660 genes)**



**C. Genes expressed in mMyos before and after Oocyte-NT
(7892 genes)**

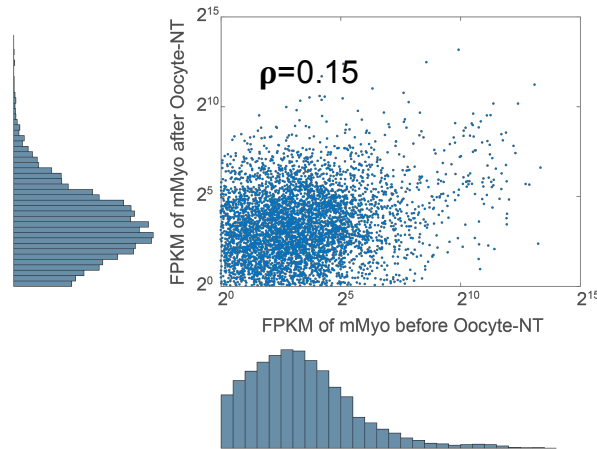


Figure 5.4.2 Genes, expressed in both donor cells and reprogrammed transcriptomes, are reprogrammed greatly by oocyte factors.

Figure 5.4.3

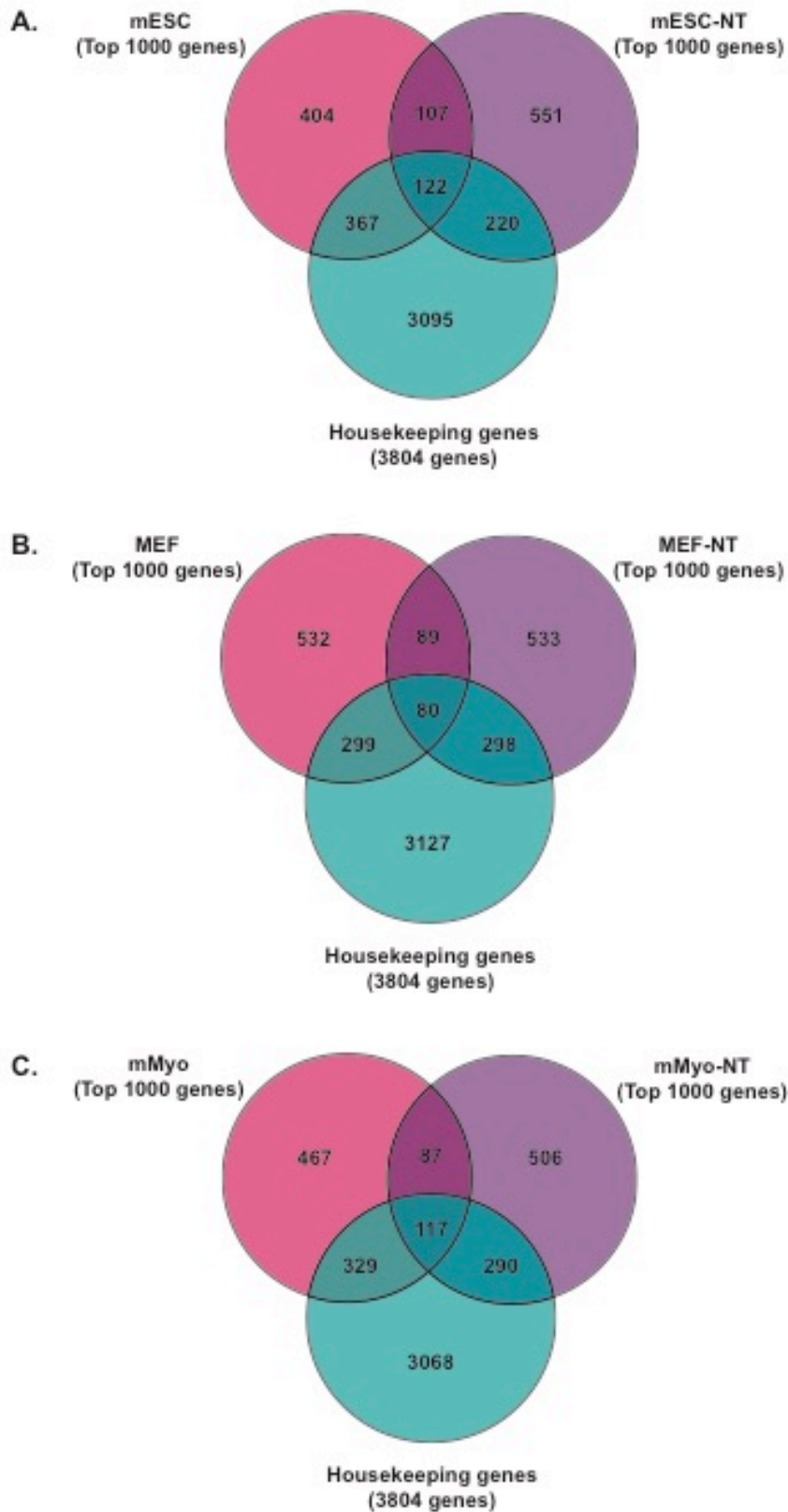


Figure 5.4.3 More than 2/3 of top 1000 genes expressed highly in donor cells are down-regulated by oocyte factors to be less than the top 1000 genes of reprogrammed transcriptomes of mESCs (A), MEFs (B) and mMyos (C).

1/3 to 1/2 of top 1000 genes for both before and after Oocyte-NT are housekeeping genes.

5.4.2 Some genes of mESCs, MEFs and mMyos are silent during SCNR by *Xenopus* oocytes

For genes silent after Oocyte-NT, the functions of these genes are usually cell-type specific and not needed for reprogrammed transcriptomes. The functions of silenced genes in mESCs, MEFs and mMyos are analyzed via Gene ontology and KEGG pathway and detailed descriptions are in Appendix (Appendix V, page 331). In addition to that, the functions of reprogrammed transcriptomes will be discussed further in 5.5.

5.4.3 Summary

In 5.3, it has been shown that oocytes can reprogram the transcriptomes of mESCs, MEFs and mMyos to a steady state within 2 days after Oocyte-NT and the majority of genes are regulated to a steady level by oocyte factors.

Through Venn diagrams, I see that 2500-5000 genes of mESCs, MEFs and mMyos are silenced and 4500-7000 genes in reprogrammed transcriptomes are activated from a silent state by oocyte factors (Figure 5.4.1). For the shared genes of transcriptomes before and after Oocyte-NT, these are regulated by oocyte factors to the level for the oocyte-determined cell identity (Figure 5.4.2). More than 2/3 of top 1000 genes expressed highly in donor cells are down-regulated by oocyte factors to be less than the top 1000 genes of reprogrammed transcriptomes. These genes are probably either repressed actively by oocyte factors or remain active but less active than genes that are more actively promoted by oocyte factors (Figure 5.4.3).

The genes silenced (32% in mESCs, 23% in MEFs and 37% in mMyos) by oocyte factors are cell-type specific to the functions of donor cells and these functional genes of donor cells are probably no longer needed for the oocyte-determined cell identity (Appendix V, page 331).

All in all, oocyte factors drive a large-scale change of transcriptomes of mESCs, MEFs and mMyos after Oocyte-NT and lead the transcriptomes of donor cells to an oocyte-determined cell identity within 48 hours.

5.5 Distinguish between reprogrammable genes and cell-type specific genes among mESC-NT, MEF-NT and mMyo-NT and the extent to which oocytes activate silent genes

Since the transcriptomes represent the identities of different cell types, the similarity and difference of transcriptomes shows the common part and the cell specific part of different cell types. It has been shown in previous sections (Section 5.2, 5.3 and 5.4) that oocytes can reprogram transcriptomes of mESCs, MEFs and mMyos to a highly similar and steady state at Day 2 after Oocyte-NT.

During SCNR by oocytes, thousands of genes in donor cells can be silenced or activated by oocyte factors and the expression of transcribed genes are regulated by oocyte factors to an oocyte-determined level. Therefore, it seems that oocytes are so competent that they can reprogram all cell types successfully but the efficiency of successful SCNR is actually low due to the existence of SCNR resistant genes. In this section, the aim is to find the cell-type specific genes that contain SCNR resistant genes and investigate the extent to which oocytes activate silent genes in the donor cells to the oocyte-steady state.

5.5.1 Reprogrammed transcriptomes of mESCs, MEFs and mMyos contain reprogrammable genes and cell-type specific genes

To look deeper into the difference between reprogrammed transcriptomes among mESCs, MEFs and mMyos, a Venn diagram is used to separate the

newly synthesized transcripts of mESC-NT, MEF-NT and mMyo-NT into two types (Figure 5.5.1).

The first type contains 11488 newly synthesized transcripts, which are expressed in all three reprogrammed transcriptomes of mESCs, MEFs and mMyos, and these genes are called reprogrammable genes (intersection of mESC-NT, MEF-NT and mMyo-NT, Figure 5.5.1).

The second type contains 10225 newly synthesized transcripts, which are expressed (FPKM>0) only in one or two reprogrammed transcriptomes and are not expressed (FPKM=0) in at least one reprogrammed transcriptomes, and they are called cell-type specific genes (Figure 5.5.1). Therefore, 2049 genes are expressed in mESC-NT and MEF-NT but not in mMyo-NT; 1815 genes are expressed in MEF-NT and mMyo-NT; 582 genes are expressed in mMyo-NT and mESC-NT (Figure 5.5.1). Moreover, 1634 genes are only expressed in mESC-NT; 3058 genes are only expressed in MEF-NT; 1087 genes are only expressed in mMyo-NT (Figure 5.5.1).

When analyzing the correlation of these newly synthesized transcripts of reprogrammed transcriptomes via Pearson correlation coefficient, the expression level of reprogrammable genes are linearly correlated between MEF-NT and mESC-NT ($\rho=0.96$), mMyo-NT and mESC-NT ($\rho=0.82$) and MEF-NT and mMyo-NT ($\rho=0.81$) (Figure 5.5.2). Notably, while comparing the expression level of cell-type specific genes, the values (ρ) of Pearson correlation coefficient between reprogrammed transcriptomes are lower than

the values for the reprogrammable genes between reprogrammed transcriptomes (Figure 5.5.2 and 5.5.3). Therefore, the expression level of cell-type specific genes between reprogrammed transcriptomes are less linearly correlated than the expression level of reprogrammable genes while the values are 0.48, 0.35 and 0.47 for the expression level of cell-type specific genes between MEF-NT and mESC-NT, mMyo-NT and mESC-NT and MEF-NT and mMyo-NT, respectively (Figure 5.5.3).

Intriguingly, if examining the histogram for the expression level of newly synthesized transcripts, the expression level of reprogrammable genes peaks higher at FPKM=4 approximately (Figure 5.5.2) than the expression level of cell-type specific genes peaks at FPKM=2 approximately (Figure 5.5.3 and 5.5.4). It suggests that the expression level of cell-type specific genes is lower because these genes are less needed for oocyte-determined cell identity or they are activated from a silent state by oocyte factors.

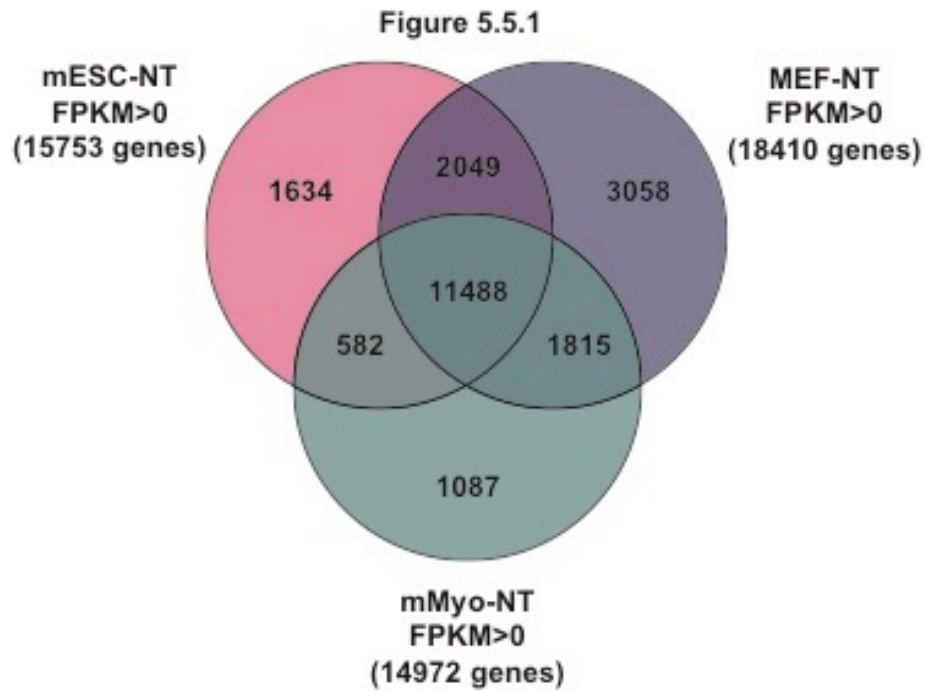


Figure 5.5.1 11488 reprogrammable genes have transcripts synthesized (FPKM>0) in reprogrammable transcriptomes of all three cell types. 10225 cell-type specific genes have newly synthesized transcripts synthesized in either one type of cells or two types of cells.

Cell-type specific genes contain genes whose expression is maintained by the transcriptional machinery of donor cells and genes that are resistant to activation by oocyte factors.

Figure 5.5.2

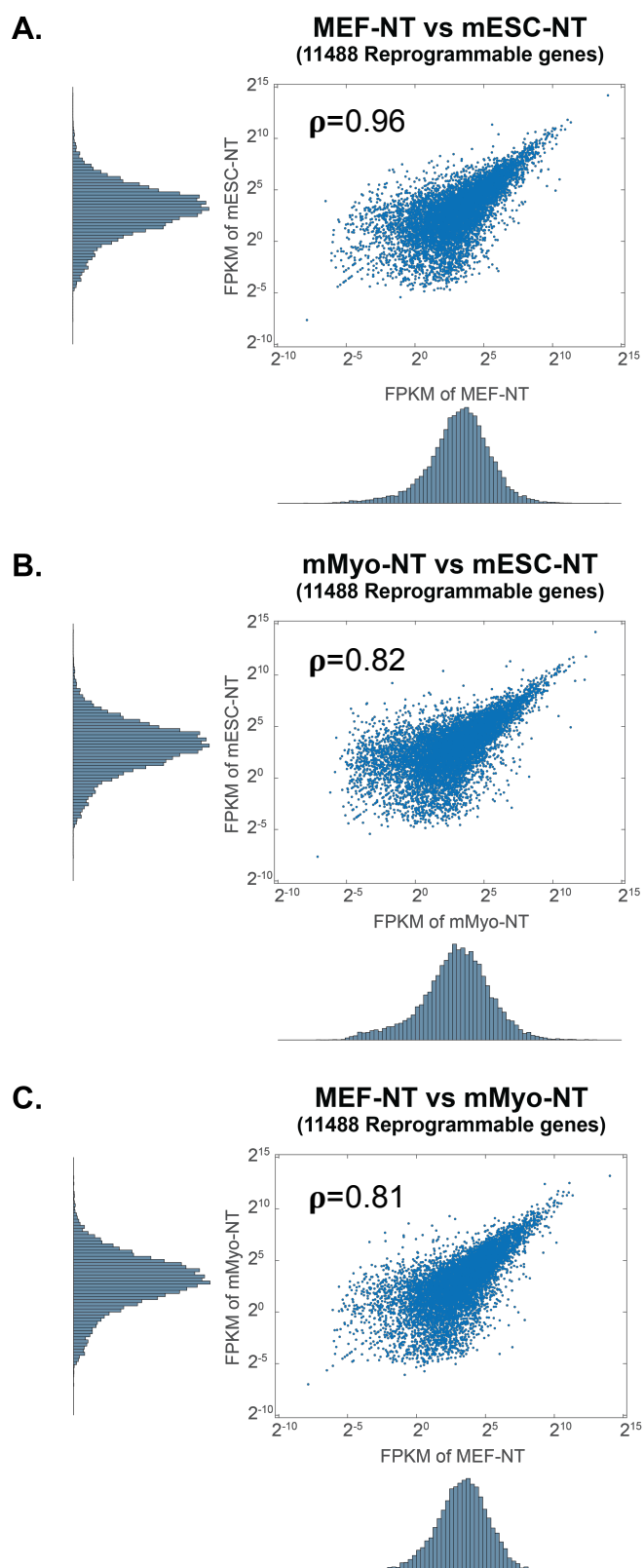


Figure 5.5.2 The expression level of reprogrammable genes, which are expressed in mESC-NT, MEF-NT and mMyo-NT, is strongly linearly-correlated between MEF-NT and mESC-NT ($\rho=0.96$), mMyo-NT and mESC-NT ($\rho=0.82$), and MEF-NT and mMyo-NT ($\rho=0.81$). The expression level of reprogrammable genes peaks at approximate FPKM=4.

Figure 5.5.3

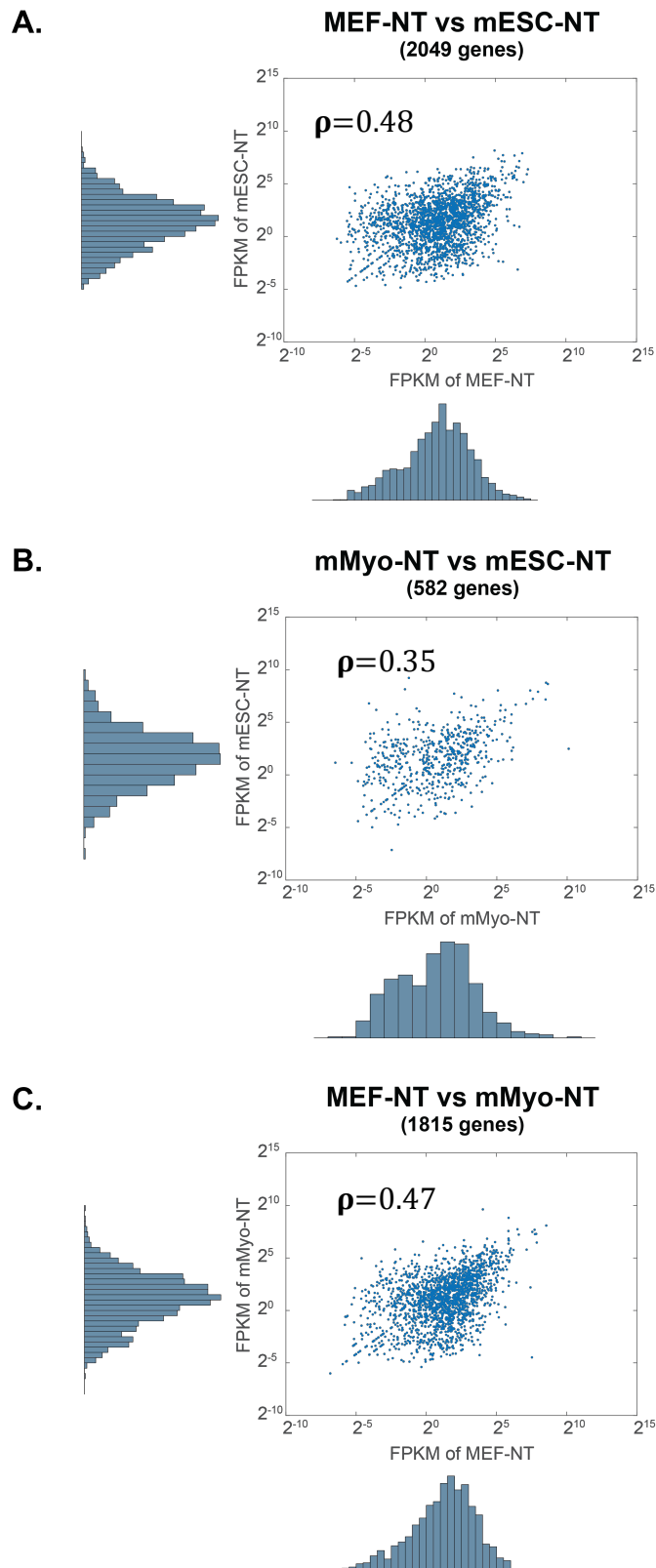


Figure 5.5.3 The cell-type specific genes expressed in two cell types are weakly linearly-correlated between MEF-NT and mESC-NT ($\rho=0.48$), mMyo-NT and mESC-NT ($\rho=0.35$), and MEF-NT and mMyo-NT ($\rho=0.47$). The expression level of cell-type specific genes peaks at approximate FPKM=2.

Figure 5.5.4

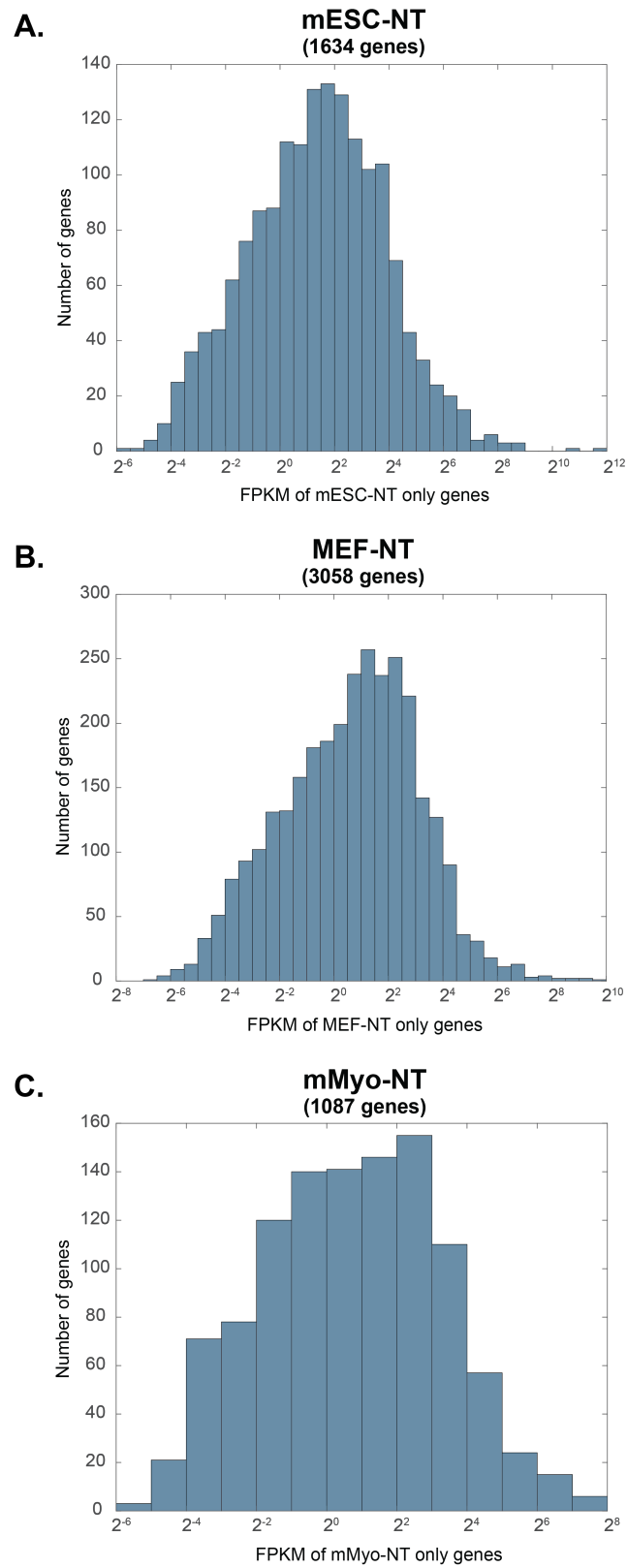


Figure 5.5.4 The cell-type specific genes, which are only expressed in one of cell types, peaks at approximate FPKM=2.

5.5.2 Reprogrammable genes possess the ability to change chromatin structures and regulate genes as the new functions of oocyte-determined cell identity

Since oocytes reprogrammed transcriptomes of donor cells to an oocyte-determined state regardless of original cell types, the functions of the oocyte-determined cell identity may be crucial to the establishment of totipotency. Therefore, reprogrammed genes are analyzed via Gene ontology and KEGG pathway for predicting the future functions of reprogrammed transcriptomes and details are noted in Appendix (Appendix VI, page 341).

Importantly, the reprogrammable transcriptomes are enriched for Signaling pathways regulating pluripotency of stem cells and this suggests that the reprogrammed transcriptomes may be primed for the establishment of pluripotency (Appendix VI, Figure 10.6.2.A, page 346).

5.5.3 Summary

In Section 5.2 and 5.3, it has been shown that oocyte factors reprogrammed transcriptomes of various cell types to an oocyte-steady state within 2 days after Oocyte-NT. Additionally, more than 2500 functional genes in transplanted mESCs, MEFs and mMyos are silenced by oocyte factors and more than 4500 genes are activated by oocyte factors (Section 5.4). Therefore, genes have new transcripts synthesized after Oocyte-NT are either activated by oocyte factors or expressed already in the donor cells.

In this section, I used Venn diagrams to categorize genes expressed in reprogrammed transcriptomes of mESCs, MEFs and mMyos as reprogrammable genes and cell-type specific genes. This shows that 11488 reprogrammable genes are expressed in reprogrammed transcriptomes of all three cell types and 10225 cell-type specific genes are expressed in reprogrammed transcriptomes of one or two cell types. Furthermore, Pearson correlation coefficients show that reprogrammable genes are reprogrammed to oocyte-steady state in all three cell types and cell-type specific genes are not.

Moreover, functions of reprogrammable genes are enriched by Gene ontology analysis and KEGG pathway. This shows that functions related to chromatin structural changes and signaling pathways regulating pluripotency of stem cells are enriched from 11488 reprogrammable genes.

All in all, reprogrammable genes are reprogrammed to oocyte-steady state by oocyte factors and have functions to make chromatin structural changes. For the cell-type specific genes, these genes are expressed in reprogrammed transcriptomes of one or two cell types with different expression level and contain one group genes whose expression is maintained by transcriptional machinery of donor cells and another group of genes that are resistant to be activated by oocyte factors.

5.6 Conclusions

It has been shown that, by comparing the range of genes regulated by oocyte factors of oocytes, transcriptomes of mESCs, MEFs and mMyos are reprogrammed to highly similar states. It suggests the expression of genes and chromatin structures are subjected to change towards the oocyte-driven state. It has also been shown that multiple genes in reprogrammed transcriptomes are involved in early embryo development, change of histone modification, chromatin remodeling and DNA conformation change. Therefore, oocyte factors have the ability to modify chromatin structure, transform the original transcriptomes and possibly promote the totipotency of donor cell nuclei.

Chapter 6 *Xenopus* klf2-HA overexpression facilitates regulation of target genes in mESCs, MEFs and mMyos during SCNR by oocytes

6.1 Background

6.1.1 Introduction

In Chapter 5, maternal factors of oocytes have been shown to reprogram mESCs, MEFs and mMyos to highly similar transcriptomes within 1 day after Oocyte-NT. It suggests the original chromatin structures of these three cell types are changed forcibly by maternal factors to a fairly similar landscape and allow maternal transcription factors to regulate target genes with modified hindrance, which resembles the change of chromatin structures and permit upcoming gene regulation to happen during early embryo development⁸².

In development, the chromatin structure of cells is subjected to change in a spatial and timely way, and to prime regulatory elements for the linked genes need to be regulated by transcription factors⁸². It has been shown that different levels of chromatin accessibility can quantitatively restrict the occupancy of transcription factors on their recognition sites⁸⁴. Additionally, genome-wide patterns of transcription factor binding shows that transcription factors can occupy their recognition sites in highly accessible chromatin regions without physical cooperative interactions⁸³. Although extracts of oocytes have been reported to epigenetically reprogram somatic cell

nuclei^{21,22}, the roles of maternal transcription factors in SCNR by oocytes are elusive¹⁴⁵. I now, in this chapter, examine the effects of TF overexpression on a wide range of genes during SCNR by oocytes. It is of particular interest to analyze oocyte components versus TF overexpression at once in the same system.

This chapter concerns a wide range of genes (15,000 to 20,000 mouse genes for differential expression analysis out of 48663 genes in the mouse reference genome) in transplanted somatic cells that maybe up- or down-regulated by oocytes with or without supplementation of xklf2-HA overexpression. In the previous Chapters 3 and 4, I have examined the effect of YF overexpression on expression of a limited number of known pluripotency gene markers during Oocyte-NT. Of the so-called Yamanaka factors, xklf2-HA has shown the greatest effect on transplanted MEF nuclei among three *Xenopus* Yamanaka factor homologs. Since the xklf2 transcripts are expressed maternally before MBT^{129,140,146} although the level of xklf2 protein counterparts are unclear^{141,142,147,148}, xklf2 may be crucial in early development and in building totipotency during SCNR by eggs.

Therefore, I investigate the effect of xklf2-HA overexpression on gene regulation during SCNR by oocytes (Section 6.2). Is the regulation of downstream genes by xklf2-HA overexpression different in various cell types during SCNR by oocytes (Section 6.3)? If xklf2 proteins maternally exist in *Xenopus* oocytes, what are the functions of xklf2 downstream genes (Section

6.4)? What are the different effects of oocyte factors and xklf2-HA overexpression on gene regulation during SCNR by oocytes (Section 6.5)?

6.1.2 Experimental design

For investigating the change after Oocyte-NT and disregarding the carried over transcripts in donor cells and *Xenopus* oocytes, BrUTP injection was performed after Oocyte-NT and the newly synthesized transcripts were incorporated with BrUTP. Before RNA-seq library preparation, the amount of BrUTP-incorporated transcripts was enriched by anti-BrUTP antibody pull-down.

The time-dependent effect of xklf2-HA overexpression on MEF-NT was evaluated by comparing oocyte samples of control and xklf2-HA groups at Day 1 and Day 2 after Oocyte-NT (Figure 6.1.1.A). The cell-type specific effects of xklf2-HA overexpression on reprogrammed transcriptomes of mESCs, MEFs and mMyos (mESC-NT, MEF-NT and mMyo-NT) were evaluated by comparing oocyte samples of control and xklf2-HA groups at Day 2 after Oocyte-NT (Figure 6.1.1.B)

At the beginning, xklf2-HA proteins were overexpressed in xklf2-HA groups by injecting xklf2-HA mRNA (9.2 ng, 1ug/ul) into oocytes 1 day before Oocyte-NT (Figure 6.1.1). For the control groups, no xklf2-HA mRNA was injected at this point. During Oocyte-NT, SLO-permeablized cells (300-500 per oocyte) were injected into GVs of oocytes in the presence or absence of xklf2-HA proteins (Figure 6.1.1). In this chapter, MEFs (mouse embryonic fibroblast cell line,

sixiFM) were used for the time-dependent effect of xklf2-HA overexpression (Figure 6.1.1.A). For the cell-type specific effects of xklf2-HA overexpression, mESCs (mouse embryonic stem cell line, B10 CHD4), sixiFM MEFs and mMyo (mouse myoblast cell line, C2C12) were tested for Oocyte-NT (Figure 6.1.1.B).

Two hours after Oocyte-NT, BrUTP (4.6 nM, 100 mM) was injected into the cytoplasm of oocytes and Oocyte-NT samples were incubated at 18°C (Figure 6.1.1). After one or two days of incubation in injected cultured oocytes, RNA was extracted, BrUTP RNA was selected, and samples were prepared for RNA-seq following cDNA synthesis from newly synthesized transcripts.

The RNA-seq data was first analyzed by Angela Simeone for alignment, hierarchical clustering, MDS (multidimensional scaling) analysis and DE (differential expression) analysis. Then, I further analysed these data for Venn diagram, Pearson correlation coefficient, Gene ontology, KEGG pathway and TRANSFAC. The preliminary data validation of RNA-seq libraries is in Appendix VII (page 348).

Figure 6.1.1

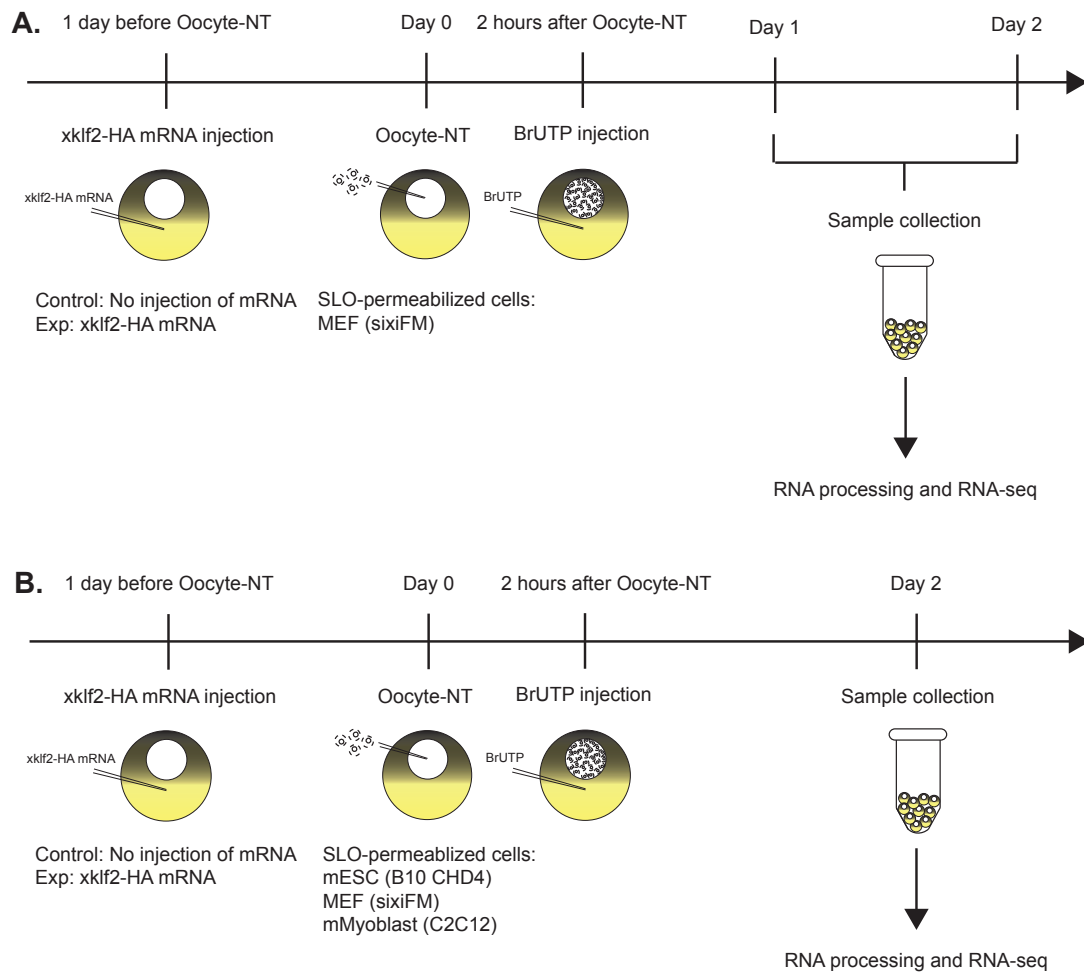


Figure 6.1.1 Sample preparations for evaluating the effect of xklf2-HA overexpression on various cell types during SCNR by oocytes

A. Time-dependent effect of xklf2-HA overexpression on MEF-NT at Day 1 and Day 2 after Oocyte-NT

B. Effects of xklf2-HA overexpression on mESC-NT, MEF-NT and mMyo-NT at Day 2 after Oocyte-NT

6.2 Overexpression of xklf2-HA reprograms MEF transcriptome within 1 day after Oocyte-NT

In Chapter 4, xklf2-HA overexpression has been shown to activate silent genes in MEFs, namely mOct4, mUtf1 and mEsrrb, at the beginning of Oocyte-NT and the difference of expression level of these genes between Control and xklf2-HA groups is diminished at later time points. In Chapter 5, it has been shown that maternal factors reprogram MEFs to an oocyte-steady state within 1 day after Oocyte-NT. In this section, I ask if xklf2-HA overexpression can also reprogram MEF nuclei within 1 day after Oocyte-NT.

In this section, I ask if xklf2-HA overexpression could also trigger a genome wide reprogramming of MEF within 1 day after Oocyte-NT. For that purpose, I compare control and xklf2-HA groups at Day 1 after Oocyte-NT. I also evaluate the time-dependent effect of xklf2-HA overexpression by comparing transcriptome at one and two days after Oocyte-NT.

6.2.1 xklf2-HA acts mostly as a transcriptional activator and xklf2-HA overexpression reprogram MEF transcriptome within 1 day after Oocyte-NT

To evaluate the short-term effect of xklf2-HA overexpression within 1 day after Oocyte-NT, MEF-NT of Control groups and xklf2-HA groups at Day 1 after Oocyte-NT are compared (Figure 6.2.1 and Table 6.2.1).

It shows that libraries of MEF-NT are different between Control groups and xklf2-HA groups genome wide and libraries are clustered by expression level of newly transcribed genes according to xklf2-HA treatment (Hierarchical

clustering, Figure 6.2.1.A; MDS analysis, Figure 6.2.1.B). This effect of *xklf2*-HA overexpression on transplanted MEF transcriptome is further evaluated using differential expression analysis. Out of the 19241 genes expressed after Oocyte-NT, 852 genes are differentially expressed (DE) in the presence of *xklf2*-HA one day after transplantation (*xklf2*-DE genes; $n=2$, $FDR<0.1$, Figure 6.2.1.C). Among these *xklf2*-DE genes in MEF-NT, 779 *xklf2*-DE genes are up-regulated ($\log_2FC>0$, $FDR<0.1$) and 73 *xklf2*-DE genes are down-regulated ($\log_2FC<0$, $FDR<0.1$) by *xklf2*-HA overexpression at Day 1 after Oocyte-NT (Figure 6.2.1.C). Therefore, the transcriptome analysis suggests that *xklf2*-HA overexpression acts mostly as a transcriptional activator during the first day after transplantation of MEF to *Xenopus* oocytes.

The distribution of 19241 analysed genes in DE analysis among different ranges of \log_2FC is further examined regardless of FDR values (Table 6.2.1) and focusing instead on the range of \log_2FC values observed. 5817 genes are up-regulated with \log_2FC more than 1 and 2659 genes are down-regulated with \log_2FC less than -1 by *xklf2*-HA overexpression (Table 6.2.1). *mKlf2*, *mSox2*, *mSall4* and *mOct4* are up-regulated with $1<\log_2FC<2$, *mUtf1* is up-regulated with $3<\log_2FC<4$ and *mEsrrb* is up-regulated with $6<\log_2FC<7$ by *xklf2*-HA overexpression at Day 1 after Oocyte-NT (BrUTP pull-down, RNA-seq, Table 6.2.1) while *mSall4* and *mOct4* have been shown to be up-regulated by *xklf2*-HA overexpression at Day 1 after Oocyte-NT (QPCR, Figure 4.2.1 in Chapter 4, page 110). Among these tested genes, only *mOct4* is shown to be up-regulated by *xklf2*-HA overexpression by both BrUTP pull-down plus RNA-seq and QPCR. Since *mOct4*, *mUtf1* and *mEsrrb* in MEFs are

silent in some of Oocyte-NT samples at Day 1 after Oocyte-NT (FPKM=0), the inconsistent up-regulation of genes, detected by QPCR and RNA-seq plus BrUTP pulldown, suggests the variable activation of silent gene by xklf2-HA overexpression at Day 1 after Oocyte-NT.

Figure 6.2.1

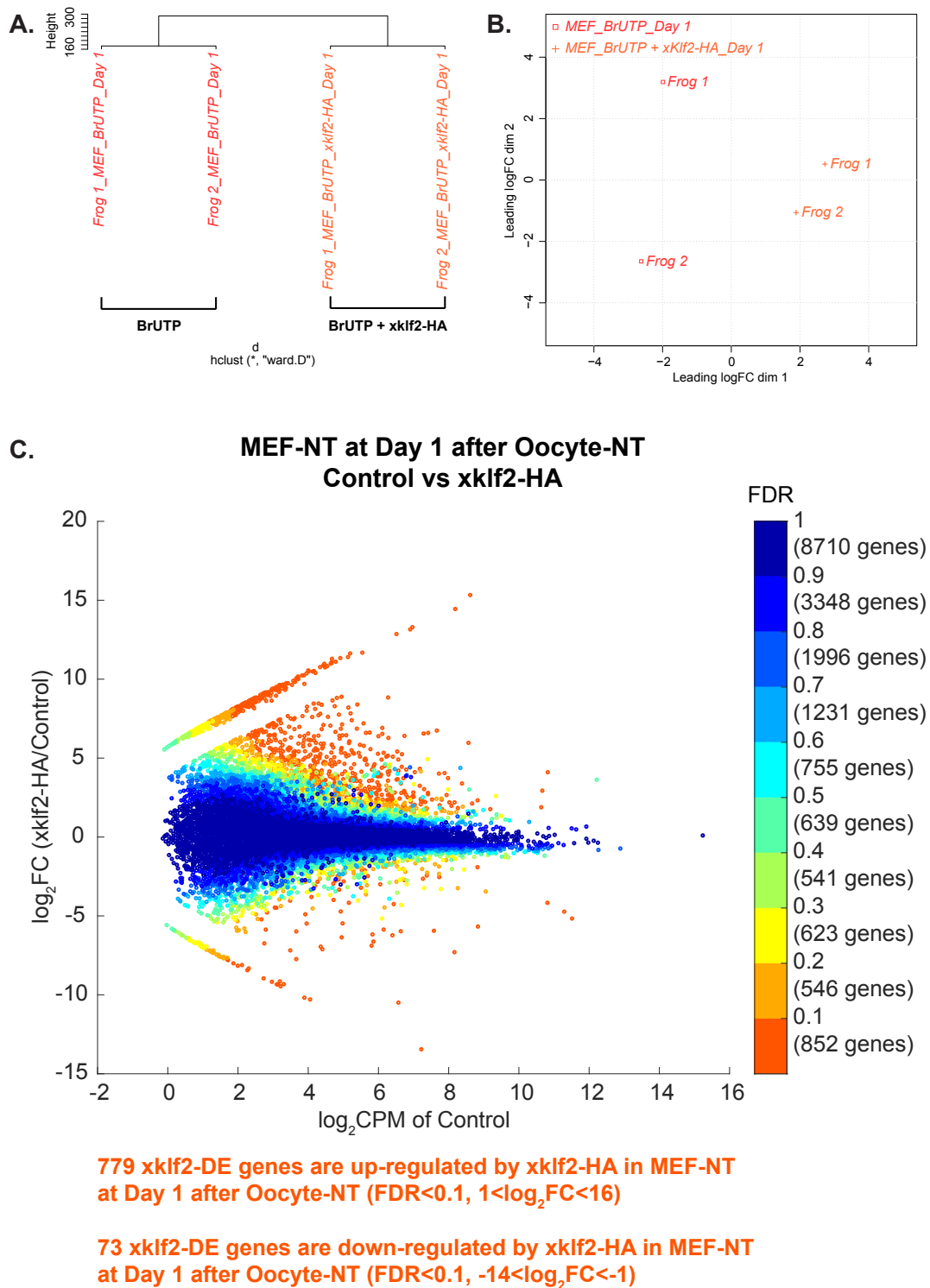


Figure 6.2.1 xklf2-HA acts mostly as a transcriptional activator at Day 1 after Oocyte-NT.
(A-B) Hierarchical clustering (A) and MDS analysis (B) show the difference between samples in the absence and presence of xklf2-HA overexpression.
(C) 779 xklf2-DE genes (FDR<0.1) are up-regulated by 2-fold to 32768-fold and 73 xklf2-DE genes in MEF-NT are down-regulated by 2-fold to 16384-fold by xklf2-HA overexpression at Day 1 after Oocyte-NT. No cut-off is applied as the threshold of log₂FC value.

Table 6.2.1

log ₂ FC	Number of genes	% of analysed genes	Pluripotency genes*
15~16	1	0.01%	
14~15	1	0.01%	
13~14	2	0.01%	
12~13	1	0.01%	
11~12	7	0.04%	
10~11	20	0.10%	
9~10	55	0.29%	
8~9	165	0.86%	
7~8	269	1.40%	
6~7	257	1.34%	mEsrrb [^]
5~6	215	1.12%	
4~5	354	1.84%	
3~4	782	4.06%	mUtf1 [^]
2~3	1311	6.81%	
1~2	2377	12.35%	mSall4 ⁺ , mOct4 ⁺⁺ , mSox2 [^] , mKlf2 [^]
0~1	4751	24.69%	mKlf4 [^]
-1~0	6014	31.26%	
-2~-1	1703	8.85%	
-3~-2	501	2.60%	
-4~-3	234	1.22%	
-5~-4	94	0.49%	
-6~-5	35	0.18%	
-7~-6	43	0.22%	
-8~-7	31	0.16%	
-9~-8	6	0.03%	
-10~-9	8	0.04%	
-11~-10	3	0.02%	
-12~-11	0	0.00%	
-13~-12	0	0.00%	
-14~-13	1	0.01%	

Total 19241 genes

log₂FC>1
5817 genes

log₂FC<-1
2659 genes

Table 6.2.1 Distribution of newly synthesized genes in MEF-NT among ranges of log₂FC (Control/xklf2-HA), judged by DE analysis, under the treatment of xklf2-HA overexpression at Day 1 after Oocyte-NT was shown. The regulation of pluripotency genes by xklf2-HA overexpression at Day 1 after Oocyte-NT is compared between DE analysis and QPCR.

* All the listed pluripotency genes are recognized as non-DE genes (FDR>0.1).

+ mSall4 is up-regulated by xklf2-HA overexpression with 1<log₂FC<2 at Day 1 after Oocyte-NT, examined by QPCR (Figure 4.2.1 in Chapter 4).

++ mOct4 is up-regulated by xklf2-HA overexpression with log₂FC>2 at Day 1 after Oocyte-NT, examined by QPCR (Figure 4.2.1 in Chapter 4).

[^] mSox2, mKlf4, mUtf1, mEsrrb and mKlf2 in MEF-NT is not regulated by xklf2-HA with log₂FC<1 at Day 1 after Oocyte-NT, examined by QPCR (Figure 4.2.1, 4.2.2 and 4.2.3 in Chapter 4).

6.2.2 Transcriptional reprogramming of transplanted cells by xklf2-HA overexpression happens mostly within 24 hours after Oocyte-NT

To examine the time-dependent effect of xklf2-HA overexpression, MEF-NT of xklf2-HA groups at Day 1 and Day 2 are compared (Figure 6.2.2 and Table 6.2.2).

Libraries of MEF-NT under the treatment of xklf2-HA overexpression at Day 1 after Oocyte-NT are grouped together and separated from libraries at Day 2 after Oocyte-NT (Hierarchical clustering, Figure 6.2.2.A; MDS analysis, Figure 6.2.2.B). From Day 1 to Day 2 after Oocyte-NT in the presence of xklf2-HA, 133 DE genes are differentially expressed (DE genes; $n=2$, $FDR<0.1$, Figure 6.2.2.C). 58 of these DE genes are up-regulated ($\log_2FC>0$, $FDR<0.1$) whereas 73 DE genes are down-regulated ($\log_2FC<0$, $FDR<0.1$) by xklf2-HA overexpression from Day 1 to Day 2 after Oocyte-NT (Figure 6.2.2.C).

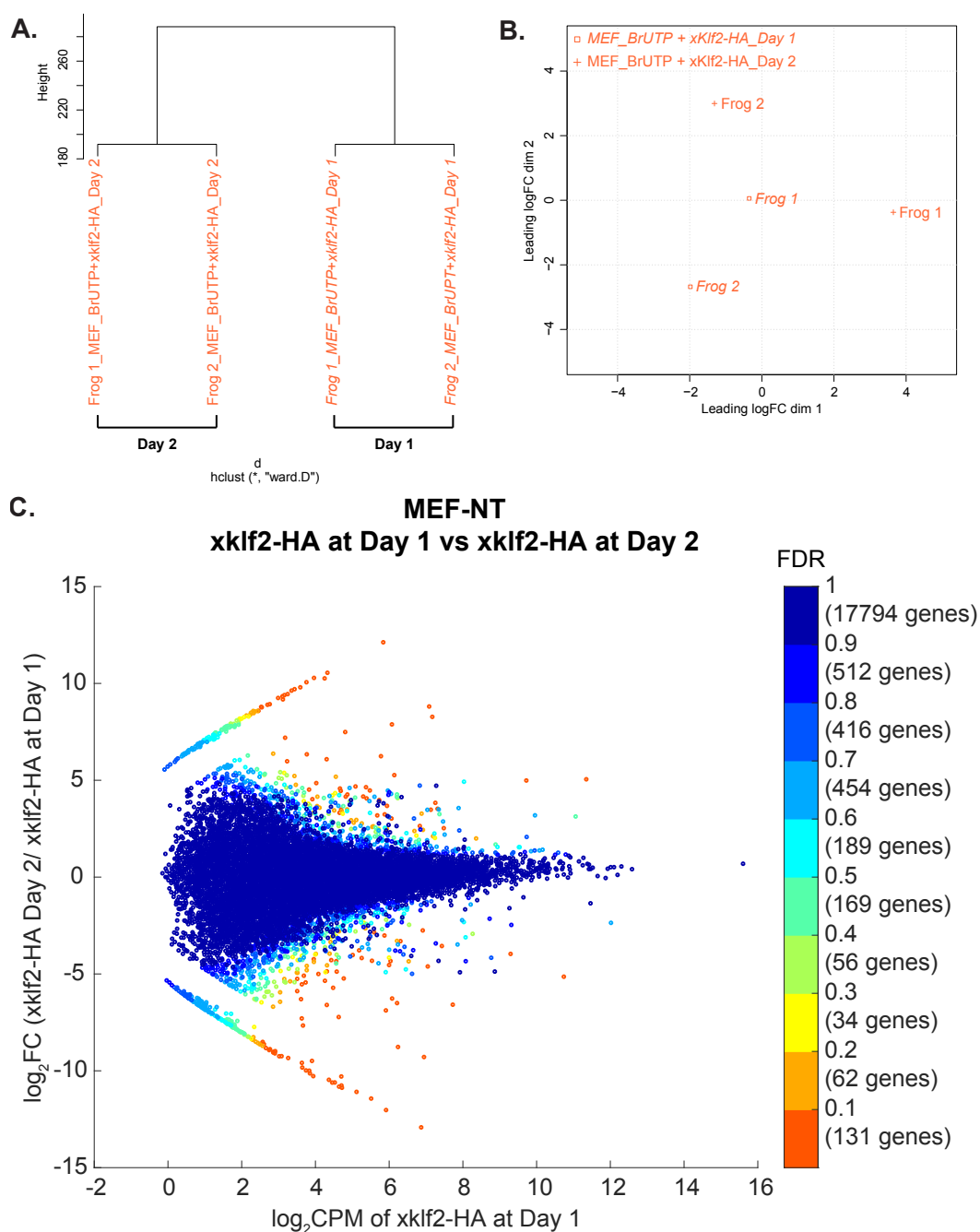
Therefore, there are much less genes regulated by xklf2-HA overexpression from Day 1 to Day 2 after Oocyte-NT (133 genes, Figure 6.2.2.C) than genes regulated by xklf2-HA at Day 1 after Oocyte-NT (852 genes, Figure 6.2.1.C). It suggests that the bulk of xklf2-HA overexpression effect on gene expression happens within 24 hours after Oocyte-NT and that few significant changes happen from Day 1 to Day 2 after Oocyte-NT.

Interestingly, the major effect of xklf2-HA overexpression on gene expression from day 1 to day 2 after Oocyte-NT is toward down-regulation of genes (73 out of 133 DE genes, Figure 6.2.2.C) whereas xklf2-HA overexpression

mostly up-regulates gene expression within the first day after Oocyte-NT (799 out of 852 *xklf2*-DE genes, Figure 6.2.1.C). Since *xklf2*-HA majorly acts as a transcriptional activator at Day 1 after Oocyte-NT (Figure 6.2.1.C), the repression of genes from Day 1 to Day 2 after Oocyte-NT in the presence of *xklf2*-HA overexpression may suggest the secondary events triggered by downstream genes of *xklf2*-HA.

Moreover, the distribution of analyzed genes located at different ranges of \log_2FC also shows that more genes are down-regulated than genes up-regulated by *xklf2*-HA overexpression from Day 1 to Day 2 after Oocyte-NT when there are 3238 genes up-regulated by *xklf2*-HA overexpression with $\log_2FC > 1$ and there are 4461 genes down-regulated by *xklf2*-HA overexpression with $\log_2FC < -1$ (Table 6.2.2). It suggests that *xklf2*-HA overexpression contribute more to up-regulate genes robustly within 1 day after Oocyte-NT but the up-regulation of some *xklf2*-HA downstream genes leads to the following down-regulation of other downstream genes from Day 1 to Day 2 after Oocyte-NT. The involvement of *xklf2*-HA downstream genes in gene regulation during SCNR by oocytes were further discussed in Appendix VIII (page 353).

Figure 6.2.2



58 DE genes are up-regulated by xklf2-HA in MEF-NT from Day 1 to Day 2 after Oocyte-NT (FDR<0.1, $2 < \log_2FC < 13$).

73 DE genes are down-regulated by xklf2-HA in MEF-NT from Day 1 to Day 2 after Oocyte-NT (FDR<0.1, $-13 < \log_2FC < -2$).

Figure 6.2.2 xklf2-HA overexpression reprogrammed MEFs within one day after Oocyte-NT.

(A-B) Hierarchical clustering (A) and MDS analysis (B) show the difference between samples collected at Day 1 and Day 2 after Oocyte-NT in the presence of xklf2-HA overexpression.

(C) 131 DE genes (FDR<0.1) are regulated by xklf2-HA overexpression from Day 1 to Day 2 after Oocyte-NT. No cut-off is applied as the threshold of \log_2FC value.

Table 6.2.2

log ₂ FC	Number of genes	% of analysed genes	Pluripotency genes*	
12~13	1	0.01%		log ₂ FC>1 3238 genes
11~12	0	0.00%		
10~11	4	0.02%		
9~10	9	0.05%		
8~9	25	0.13%		
7~8	49	0.25%		
6~7	43	0.22%		
5~6	57	0.29%		
4~5	157	0.79%		
3~4	405	2.04%		
2~3	674	3.40%	mUtf1+	
1~2	1814	9.15%	mSox2+	
0~1	6030	30.43%	mEsrrb+, mKlf4+, mSall4++, mOct4++	log ₂ FC<-1 4461 genes
-1~0	6087	30.72%	mKlf2^	
-2~-1	2252	11.36%		
-3~-2	1025	5.17%		
-4~-3	504	2.54%		
-5~-4	204	1.03%		
-6~-5	122	0.62%		
-7~-6	157	0.79%		
-8~-7	131	0.66%		
-9~-8	39	0.20%		
-10~-9	12	0.06%		
-11~-10	11	0.06%		
-12~-11	2	0.01%		
-13~-12	2	0.01%		

Total 19816 genes

Table 6.2.2 Distribution of newly synthesized genes in MEF-NT among ranges of log₂FC (xklf2-HA at Day 2/xklf2-HA at Day 1), judged by DE analysis, under the treatment of xklf2-HA overexpression from Day 1 to Day 2 after Oocyte-NT. The regulation of pluripotency genes by xklf2-HA overexpression from Day 1 to Day 2 after Oocyte-NT is compared between DE analysis and QPCR.

* All the listed pluripotency genes are recognized as non-DE genes (FDR>0.1).

+ mUtf1, mSox2, mEsrrb, mKlf4 is up-regulated by xklf2-HA with 1<log₂FC<2 from Day 1 to Day 2 after Oocyte-NT, examined by QPCR (Figure 4.2.1 and 4.2.2, page 110 and 111).

++ mSall4 and mOct4 is up-regulated by xklf2-HA with log₂FC>2 from Day 1 to Day 2 after Oocyte-NT, examined by QPCR (Figure 4.2.1, page 110).

^ mKlf2 is not regulated by xklf2-HA overexpression with log₂FC<1 from Day 1 to Day 2 after Oocyte-NT, examined by QPCR (Figure 4.2.3 , page 112).

6.2.3 Summary

In Chapter 5, it has been shown that maternal factors of oocytes reprogram MEFs to an oocyte-steady state within 1 day after Oocyte-NT and the speed of SCNR by oocytes can be varied due to the batch effects of oocytes. Likewise, *xklf2*-HA overexpression is shown to regulate genes in MEF-NT during SCNR by oocytes and reprogram MEFs to a *xklf2*-oocyte state within 1 day after Oocyte-NT.

When comparing genes regulated by *xklf2*-HA overexpression before and after Day 1 after Oocyte-NT, it shows that different sets of downstream genes are regulated by *xklf2*-HA overexpression at different time points. Additionally, more genes are up-regulated than down-regulated by *xklf2*-HA overexpression within Day 1 after Oocyte-NT while it is opposite after Day 1 after oocyte-NT. It suggests not only maternal factors and *xklf2*-HA combinatorially regulates gene expression during SCNR by oocytes but downstream genes of *xklf2*-HA also participate in the gene regulation during SCNR by oocytes after they are up-regulated by *xklf2*-HA overexpression. This point can be further validated by cycloheximide treatment.

To conclude, *xklf2*-HA overexpression reprograms MEF-NT to an *xklf2*-oocyte state within 1 day after Oocyte-NT. Moreover, *xklf2*-HA overexpression regulates different sets of genes at different time points because downstream genes of *xklf2*-HA also participate in gene regulation during SCNR by oocytes.

6.3 Overexpression of xklf2-HA facilitates the regulation of gene expression in mESCs, MEFs and mMyos during SCNR by oocytes

In Chapter 5, maternal factors of *Xenopus* oocytes have been shown to reprogram transcriptomes of mESCs, MEFs and mMyos to an oocyte-steady state within 2 days after Oocyte-NT. Because compositions and amount of maternal factors are fixed in *Xenopus* oocytes and chromatin structures of various donor cell types are different, it suggests chromatin structures of mESCs, MEFs and mMyos are subjected to be changed by maternal factors and allow transcriptional machinery in oocytes to regulate expression of genes to an oocyte-steady level.

In this section, I would like to ask the way by which xklf2-HA overexpression regulates gene expression in mESCs, MEFs and mMyos during SCNR by oocytes while chromatin structures of various cell types are dynamically changed at the same time?

6.3.1 Overexpression of xklf2-HA regulates more than 1000 genes in mESCs, MEFs and mMyos at Day 2 after Oocyte-NT

To examine the effect of xklf2-HA overexpression on mESC-NT, MEF-NT and mMyo-NT, Control groups and xklf2-HA groups are compared at Day 2 after Oocyte-NT.

Hierarchical clustering shows that reprogrammed transcriptomes of mESCs, MEFs and mMyos are clustered by expression level of newly synthesized transcripts genome wide and libraries are grouped together based on the xklf2-HA treatment (Figure 6.3.1.A, 6.3.2.A and 6.3.3.A). Similarly, Control groups and xklf2-HA groups of mESC-NT, MEF-NT and mMyo-NT are shown to fall separately on MDS plot (Figure 6.3.1.B, 6.3.2.B and 6.3.3.B). This indicates that xklf2-HA overexpression triggers a global shift in reprogrammed transcriptomes of mESCs, MEFs and mMyos irrespective of origin of donor cells and batch effects of *Xenopus* oocytes.

In mESC-NT, 1110 xklf2-DE genes are significantly regulated by xklf2-HA overexpression at Day 2 after Oocyte-NT (n=3, FDR<0.1, Figure 6.3.1.C). Among these genes, 976 xklf2-DE genes in mESC-NT are up-regulated by more than 2-fold ($2 < FC < 4096$, FDR<0.1) and 134 xklf2-DE genes in mESC-NT are down-regulated by more than 2-fold ($-4096 < \text{Fold change} < -2$, FDR<0.1) by xklf2-HA at Day 2 after Oocyte-NT (Figure 6.3.1.C). Similarly, in MEF-NT, among 2871 xklf2-DE genes that are significantly regulated (n=4, FDR<0.1, Figure 6.3.2.C), 2044 xklf2-DE genes are up-regulated by more than 1-fold ($1 < FC < 4096$, FDR<0.1) and 827 xklf2-DE genes are down-regulated by more

than 1-fold ($-512 < FC < 1$, $FDR < 0.1$) by *xklf2*-HA overexpression (Figure 6.3.2.C). Lastly, in mMyo-NT, there are 1077 *xklf2*-DE genes regulated by *xklf2*-HA overexpression ($n=3$, $FDR < 0.1$, Figure 6.3.3): 936 *xklf2*-DE genes in mMyo-NT are up-regulated ($2 < FC < 8192$, $FDR < 0.1$) and 141 *xklf2*-DE genes are down-regulated ($-2048 < FC < -2$, $FDR < 0.1$) by *xklf2*-HA overexpression (Figure 6.3.3.C). As it is observed that more than 2/3 of *xklf2*-DE genes are up-regulated by *xklf2*-HA overexpression in mESC-NT, MEF-NT and mMyo-NT, it suggests that *xklf2*-HA mostly acts as a transcriptional activator during the time period analysed.

Overall, *xklf2*-HA overexpression affects more than 1000 genes in mESC-NT, MEF-NT and mMyo-NT at Day 2 after Oocyte-NT (Figure 6.3.1.C, 6.3.2.C and 6.3.3.C). Notably, there are more *xklf2*-DE genes identified in MEF-NT. This is because the number of biological replicates used for MEF-NT goes up to 4, rather than 3 for mESC-NT and mMyo-NT. Utilizing more biological replicates improves the identification of *xklf2*-DE genes while the number of *xklf2*-DE genes increases and more down-regulated *xklf2*-DE genes are found.

In terms of fold change, *xklf2*-HA overexpression up-regulates *xklf2*-DE genes at lower expression level with higher $\log_2 FC$ ($\log_2 FC > 0$, $FDR < 0.1$) and down-regulates *xklf2*-DE genes at lower expression level with lower $\log_2 FC$ ($\log_2 FC < 0$, $FDR < 0.1$) significantly in mESC-NT, MEF-NT and mMyo-NT (Appendix IX, page 357).

Figure 6.3.1

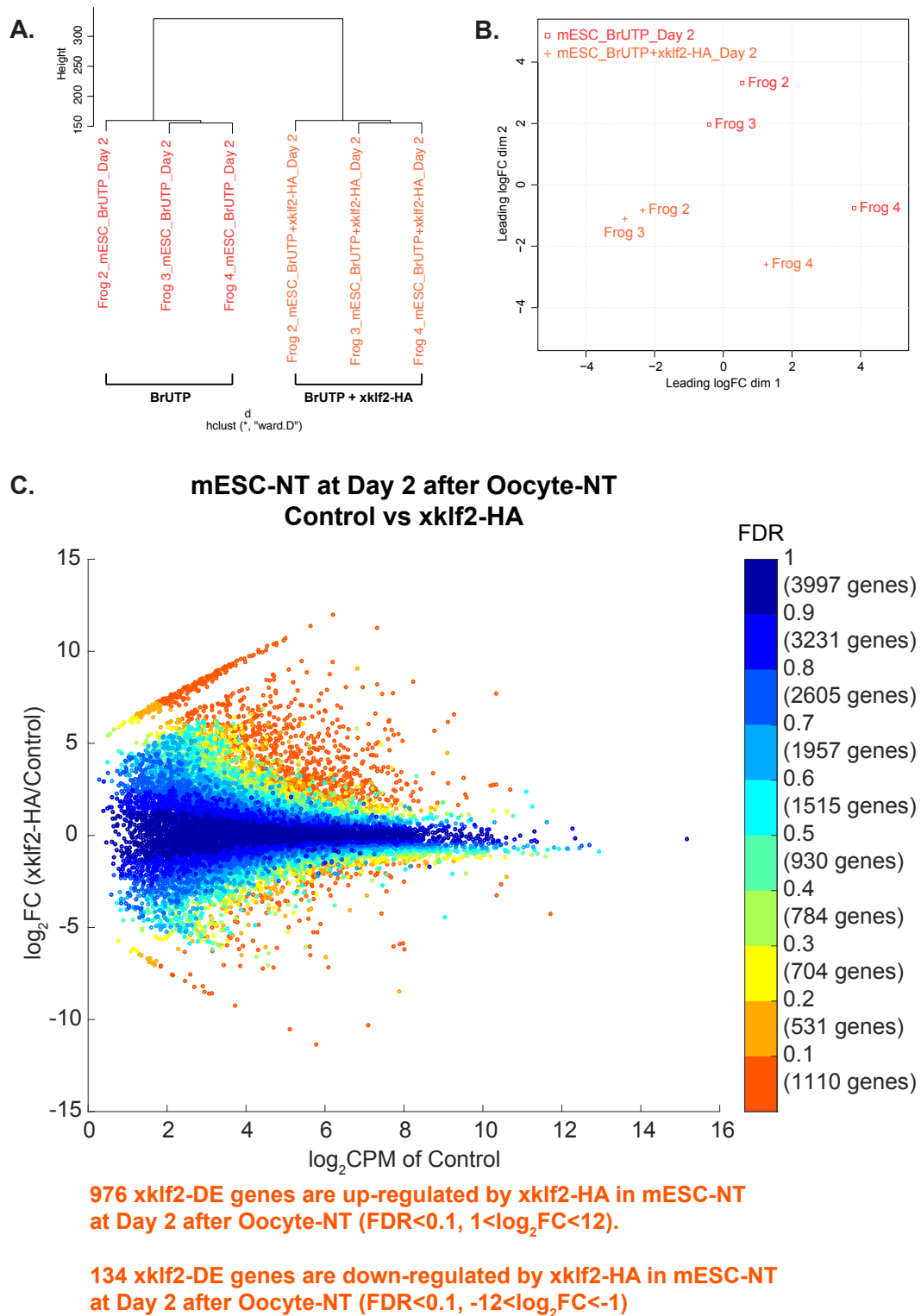


Figure 6.3.1 xklf2-HA overexpression shift transcriptomes of mESCs genome wide at Day 2 after Oocyte-NT. Hierarchical clustering (A) and MDS analysis (B) show the difference between samples in the absence (in red) and presence (in orange) of xklf2-HA overexpression. (C) 1110 xklf2-DE genes (FDR<0.1, in orange) are found by DE analysis. No cut-off is applied as the threshold of log₂FC value.

Figure 6.3.2

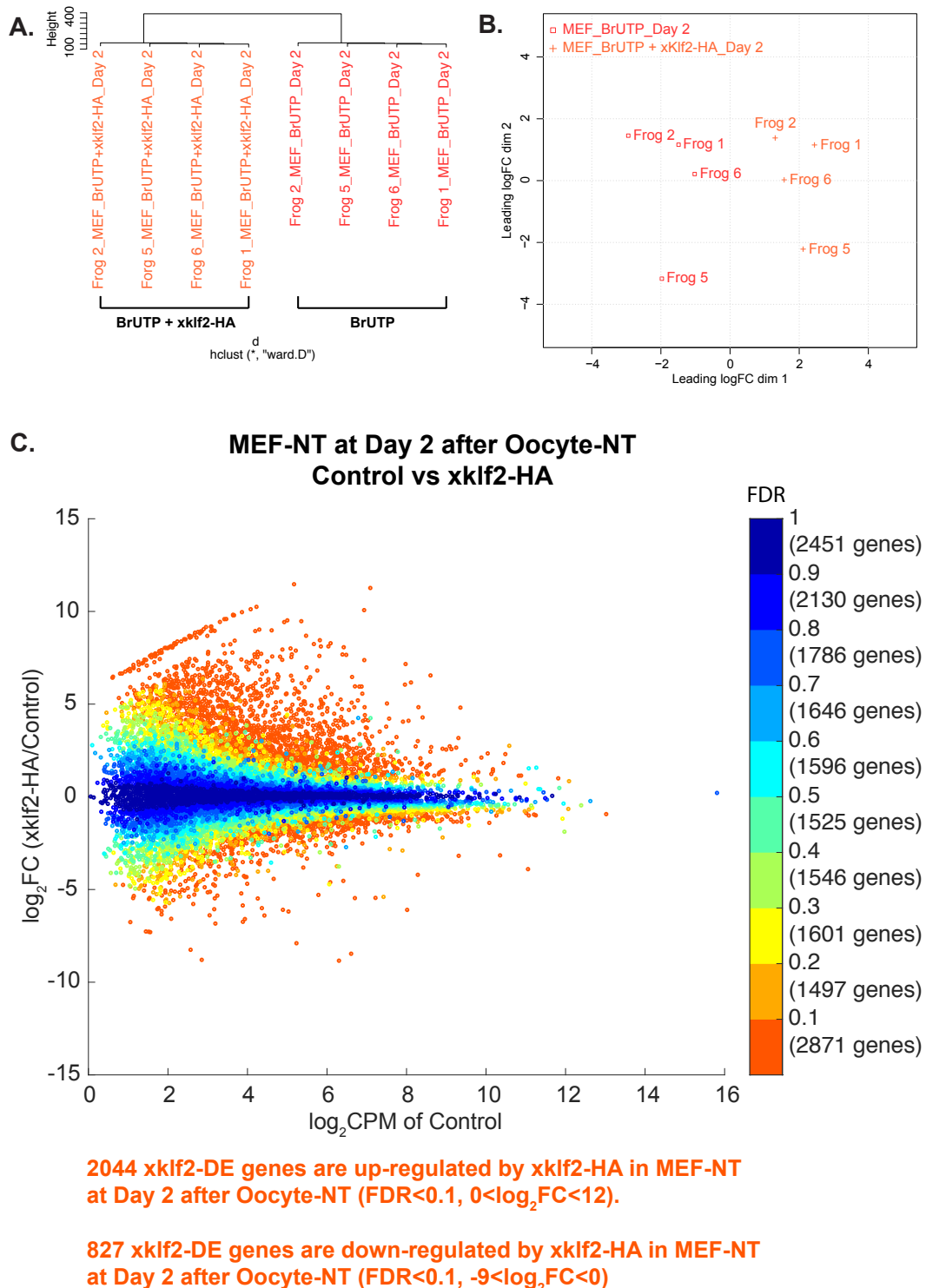


Figure 6.3.2 xklf2-HA overexpression shift transcriptomes of MEFs genome wide at Day 2 after Oocyte-NT. Hierarchical clustering (A) and MDS analysis (B) show the difference between samples in the absence (in red) and presence (in orange) of xklf2-HA overexpression. (C) 2871 xklf2-DE genes (FDR<0.1, in orange) are found by DE analysis. No cut-off is applied as the threshold of log₂FC value.

Figure 6.3.3

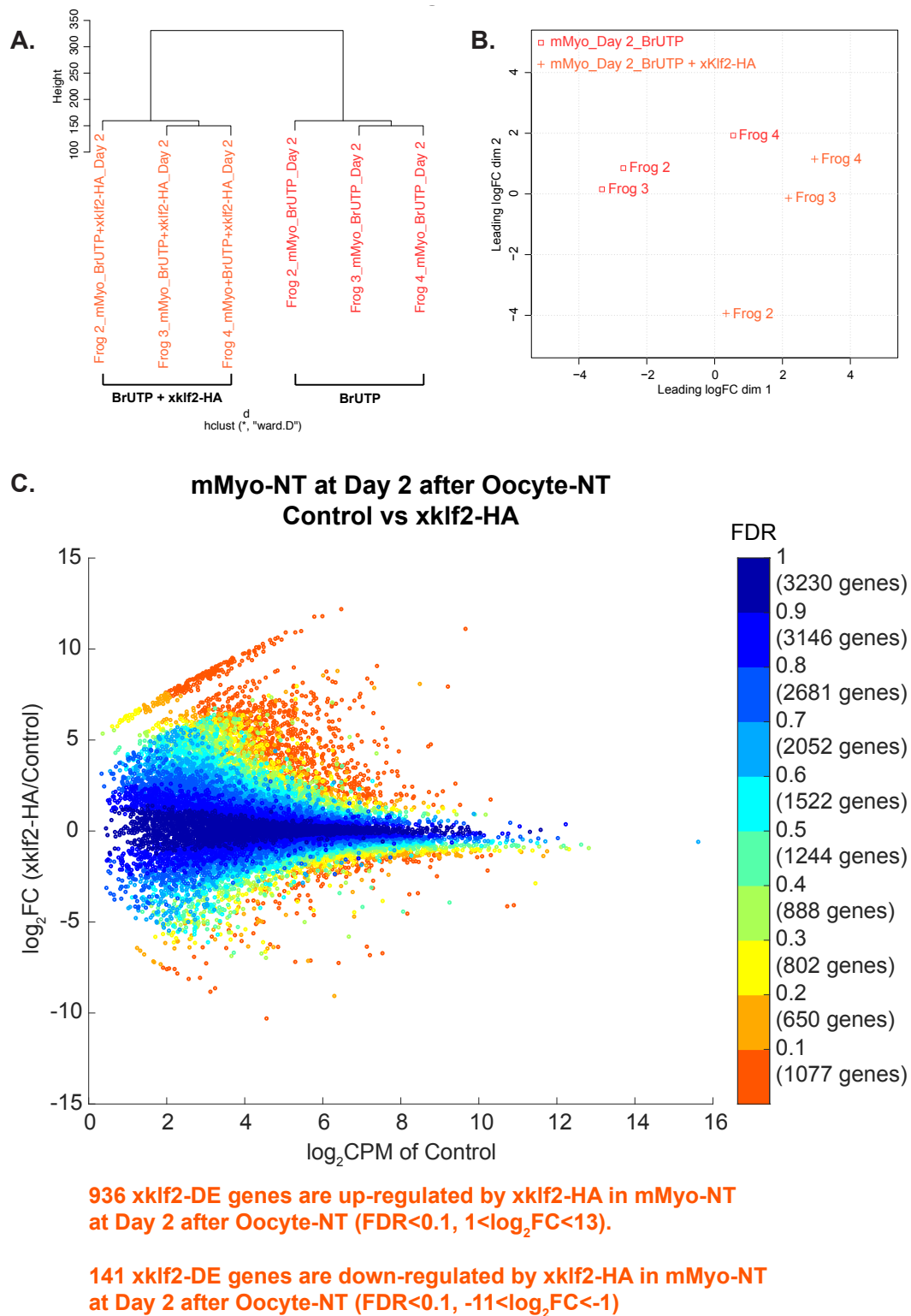


Figure 6.3.3 xklf2-HA overexpression shift transcriptomes of mMyos genome wide at Day 2 after Oocyte-NT. Hierarchical clustering (A) and MDS analysis (B) show the difference between samples in the absence (in red) and presence (in orange) of xklf2-HA overexpression. (C) 2871 xklf2-DE genes (FDR<0.1, in orange) are found by DE analysis. No cut-off is applied as the threshold of log₂FC value.

6.3.2 xklf2-HA overexpression facilitates regulation of downstream genes after Oocyte-NT in MEF-NT

To evaluate regulation of newly synthesized genes by xklf2-HA overexpression regardless of FDR, Control groups and xklf2-HA groups are compared (Table 6.3.1). In MEF-NT, there are 5100 genes up-regulated by xklf2-HA overexpression ($\log_2FC > 1$) and 3503 genes down-regulated by xklf2-HA overexpression ($\log_2FC < -1$) (Table 6.3.1). For up-regulation of genes by xklf2-HA overexpression, 30% of analyzed genes are up-regulated at Day 1 after Oocyte-NT (Table 6.2.1, page 197) and 27% are up-regulated at Day 2 after Oocyte-NT (Table 6.3.1). For down-regulation of genes by xklf2-HA overexpression, 14% of analyzed genes are down-regulated at Day 1 after Oocyte-NT (Table 6.2.1, page 197) and 19% are down-regulated at Day 2 after Oocyte-NT (Table 6.3.1).

Therefore, less genes are up-regulated by xklf2-HA overexpression at Day 2 than at Day 1 after Oocyte-NT and more genes are down-regulated by xklf2-HA overexpression at Day 2 than at Day 1 after Oocyte-NT (Table 6.2.1 and 6.3.1) because more genes are down-regulated than up-regulated by xklf2-HA overexpression from Day 1 to Day 2 after Oocyte-NT (Table 6.2.2, page 201). At Day 2 after Oocyte-NT, two pluripotency genes in MEFs, namely mOct4 and mSall4, are recognized as xklf2-DE genes ($FDR < 0.1$, Figure 6.3.2; Table 6.3.1). Since mOct4 is silent in MEFs ($FPKM = 0$, ENCODE) and at nil/low expression level in MEF-NT by chance at Day 2 after Oocyte-NT while it is not successfully activated by maternal factors ($FPKM = 0$), xklf2-HA overexpression promotes the activation of silent genes, such as mOct4.

Table 6.3.1

log ₂ FC	Number of genes	% of analysed genes	Pluripotency genes	
11~12	2	0.01%		log ₂ FC>1 5100 genes
10~11	3	0.02%		
9~10	13	0.07%		
8~9	44	0.24%		
7~8	78	0.42%		
6~7	142	0.76%		
5~6	252	1.35%		
4~5	386	2.07%		
3~4	696	3.73%	mOct4*++	
2~3	1212	6.50%	mSall4*++	
1~2	2272	12.18%	mUtf1+, mSox2*+, mEsrrb [^] , mKlf2* [^]	
0~1	4258	22.83%	mKlf4+	log ₂ FC<-1 3503 genes
-1~0	5788	31.04%		
-2~-1	2364	12.68%		
-3~-2	658	3.53%		
-4~-3	286	1.53%		
-5~-4	125	0.67%		
-6~-5	46	0.25%		
-7~-6	16	0.09%		
-8~-7	4	0.02%		
-9~-8	4	0.02%		

Total 18649 genes

Table 6.3.1 Distribution of newly synthesized genes in MEF-NT among ranges of log₂FC (Control/xklf2-HA), judged by DE analysis, under the treatment of xklf2-HA overexpression at Day 2 after Oocyte-NT and the regulation of pluripotency genes by xklf2-HA at Day 2 after Oocyte-NT are compared between DE analysis and QPCR.

* mOct4, mSall4 and mSox2 are recognized as xklf2-DE genes (FDR<0.1)

+ mUtf1, mSox2, mKlf4 in MEF-NT is up-regulated by xklf2-HA overexpression with 1<log₂FC<2 at Day 2 after Oocyte-NT, examined by QPCR (Figure 4.2.1, page 110).

++ mOct4 and mSall4 in MEF-NT is up-regulated by xklf2-HA overexpression with log₂FC>2 at Day 2 after Oocyte-NT, examined by QPCR (Figure 4.2.1 , page 110).

[^] mEsrrb and mKlf2 in MEF-NT is not regulated by xklf2-HA with log₂FC<1 at Day 2 after Oocyte-NT, examined by QPCR (Figure 4.2.2 and 4.2.3, page 111 and 112).

6.3.3 Summary

In Chapter 5, it has been shown that maternal factors of *Xenopus* oocytes reprogram transcriptomes of mESCs, MEFs and mMyos to an oocyte-steady state at Day 2 after Oocyte-NT. The fixed composition and amount of maternal factors change the chromatin structure of mESCs, MEFs and mMyos and regulate expression of genes to an oocyte-steady level.

During this dynamic SCNR process by oocytes, xklf2-HA overexpression regulated more than 1000 xklf2-DE genes in mESC-NT, MEF-NT and mMyo-NT significantly at Day 2 after Oocyte-NT. More than half of xklf2-DE genes are at nil or low expression level. While expression level of genes correlates to chromatin accessibility, xklf2-HA overexpression tends to regulate genes with regulatory elements reside in closed chromatin. Since mOct4 has shown to be silent genes in MEFs and in some MEF-NT when mOct4 is not successfully activated by maternal factors, xklf2-HA overexpression promotes the activation of silent genes.

6.4 Overexpression of xklf2-HA regulates xklf2-DE genes to the xklf2-oocyte level in mESC-NT, MEF-NT and mMyo-NT at Day 2 after Oocyte-NT

In Chapter 5, it has been shown that maternal factors reprogram transcriptomes of mESCs, MEFs and mMyos to a fairly similar state in *Xenopus* oocytes. In this section, the genome-wide regulation of downstream genes by xklf2-HA overexpression in mESCs, MEFs and mMyos is evaluated during this dynamic change of SCNR by oocytes and I would like to ask if xklf2-HA overexpression would affect different sets of downstream genes due to original chromatin structures of mESCs, MEFs and mMyos?

6.4.1 xklf2-HA overexpression regulates different sets of xklf2-DE genes in mESC-NT, MFE-NT and mMyo-NT at Day 2 after Oocyte-NT while xklf2-DE genes are either up-regulated or down-regulated by xklf2-HA overexpression

To distinguish cell-type specific effect of xklf2-HA overexpression on the regulation of downstream genes in mESC-NT, MEF-NT and mMyo-NT, up-regulated and down-regulated xklf2-DE genes in mESC-NT, MEF-NT and mMyo-NT at Day 2 after Oocyte-NT are compared via Venn diagram (Figure 6.4.1 and 6.4.2).

Among all xklf2-DE genes in mESC-NT, MEF-NT and mMyo-NT, 347 xklf2-DE genes are either up-regulated or down-regulated by xklf2-HA overexpression in all 3 cell types (Figure 6.4.1.A). 386 xklf2-DE genes are

only regulated in mESC-NT; 1876 xklf2-DE genes are only regulated in MEF-NT; 371 xklf2-DE genes are only regulated in mMyo-NT (Figure 6.4.1.A).

For up-regulated xklf2-DE genes, 333 xklf2-DE genes are up-regulated in all 3 cell types (Figure 6.4.1.B). In mESC-NT, 65% of up-regulated xklf2-DE genes are also up-regulated in MEF-NT and mMyo-NT and 35% of up-regulated xklf2-DE genes are only up-regulated in mESC-NT (Figure 6.4.1.B). In MEF-NT, 35% of up-regulated xklf2-DE genes are also up-regulated in mESC-NT and mMyo-NT and 65% of up-regulated xklf2-DE genes are only up-regulated in MEF-NT (Figure 6.4.1.B). In mMyo-NT, 66% of up-regulated xklf2-DE genes are also up-regulated in mESC-NT and MEF-NT and 34% of up-regulated xklf2-DE genes are only up-regulated in mMyo-NT (Figure 6.4.1.B). Therefore, more than 30% of up-regulated xklf2-DE genes are only up-regulated by xklf2-HA overexpression in one cell type.

For down-regulated xklf2-DE genes, 12 xklf2-DE genes are down-regulated in all 3 cell types (Figure 6.4.1.C). In mESC-NT, 32% of down-regulated xklf2-DE genes are also down-regulated in MEF-NT and mMyo-NT and 68% of down-regulated xklf2-DE genes are only down-regulated in mESC-NT (Figure 6.4.1.C). In MEF-NT, 9% of down-regulated xklf2-DE genes are also down-regulated in mESC-NT and mMyo-NT and 91% of down-regulated xklf2-DE genes are only down-regulated in MEF-NT (Figure 6.4.1.B). In mMyo-NT, 40% of down-regulated xklf2-DE genes are also down-regulated in mESC-NT and MEF-NT and 60% of up-regulated xklf2-DE genes are only up-regulated in mMyo-NT (Figure 6.4.1.C). Hence, more than 60% of down-regulated xklf2-

DE genes are only down-regulated by xklf2-HA overexpression in one cell type.

When comparing all up-regulated and down-regulated xklf2-DE genes in mESC-NT, MEF-NT and mMyo-NT together, the majority of up-regulated xklf2-DE genes in any of these three cell types are not down-regulated in other cell types (Figure 6.4.2). Only few xklf2-DE genes are up-regulated in one cell type and down-regulated in other cell types or reversely (number of xklf2-DE genes in red, Figure 6.4.2). It indicates the regulation of genes by xklf2-HA overexpression are mostly in one direction within 2 day after Oocyte-NT and xklf2-DE genes are either up-regulated or down-regulated regardless of different transplanted cell types.

Figure 6.4.1

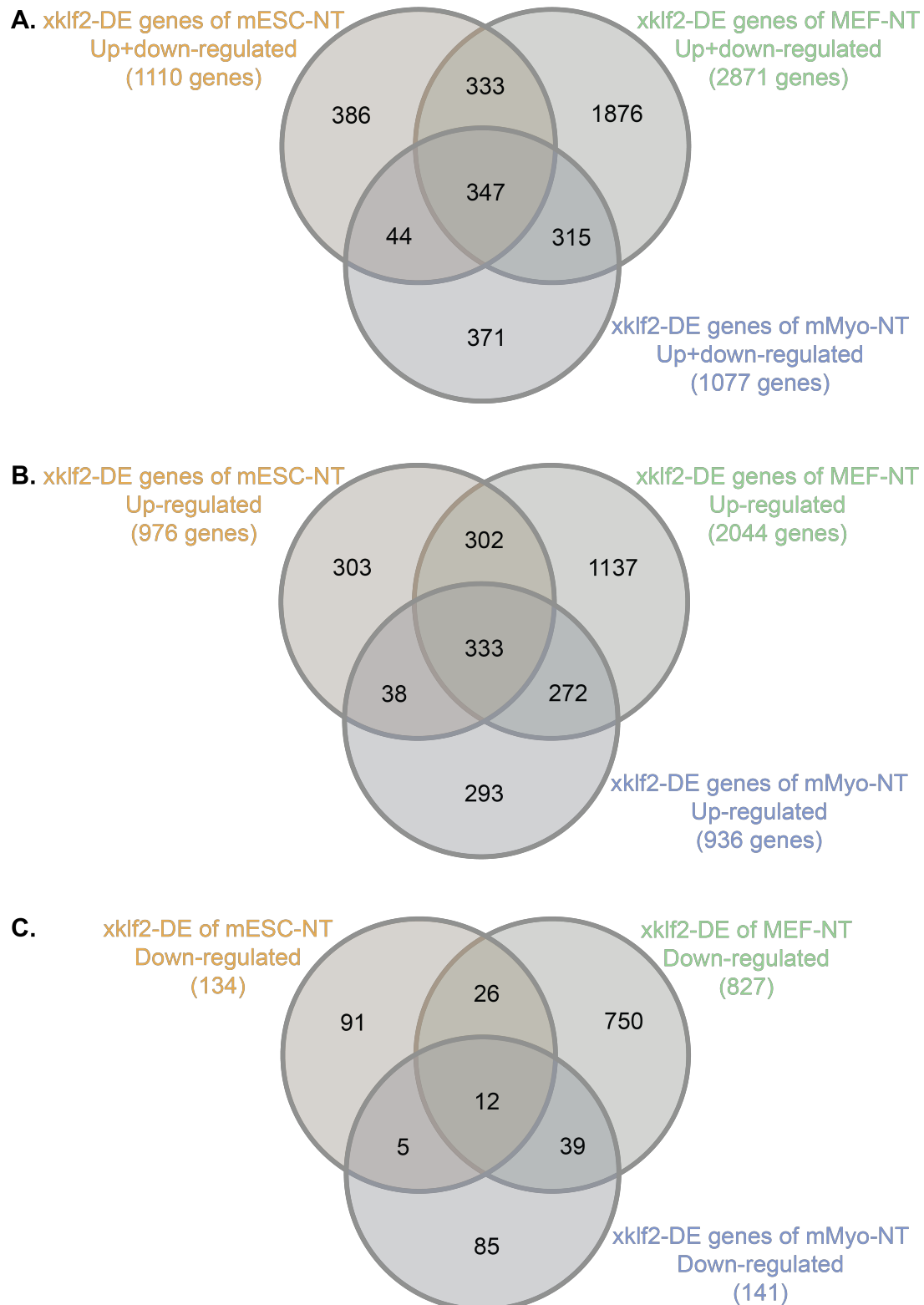


Figure 6.4.1 The majority of shared xklf2-DE genes among mESC-NT, MEF-NT and mMyo-NT (A) are contributed by up-regulated xklf2-DE genes at Day 2 after Oocyte-NT (B and C). 31%, 56% and 31% of up-regulated (B) and 68%, 91% and 60% of down-regulated (C) xklf2-DE genes are cell-type specific and only regulated in mESC-NT, MEF-NT and mMyo-NT, respectively.

Figure 6.4.2

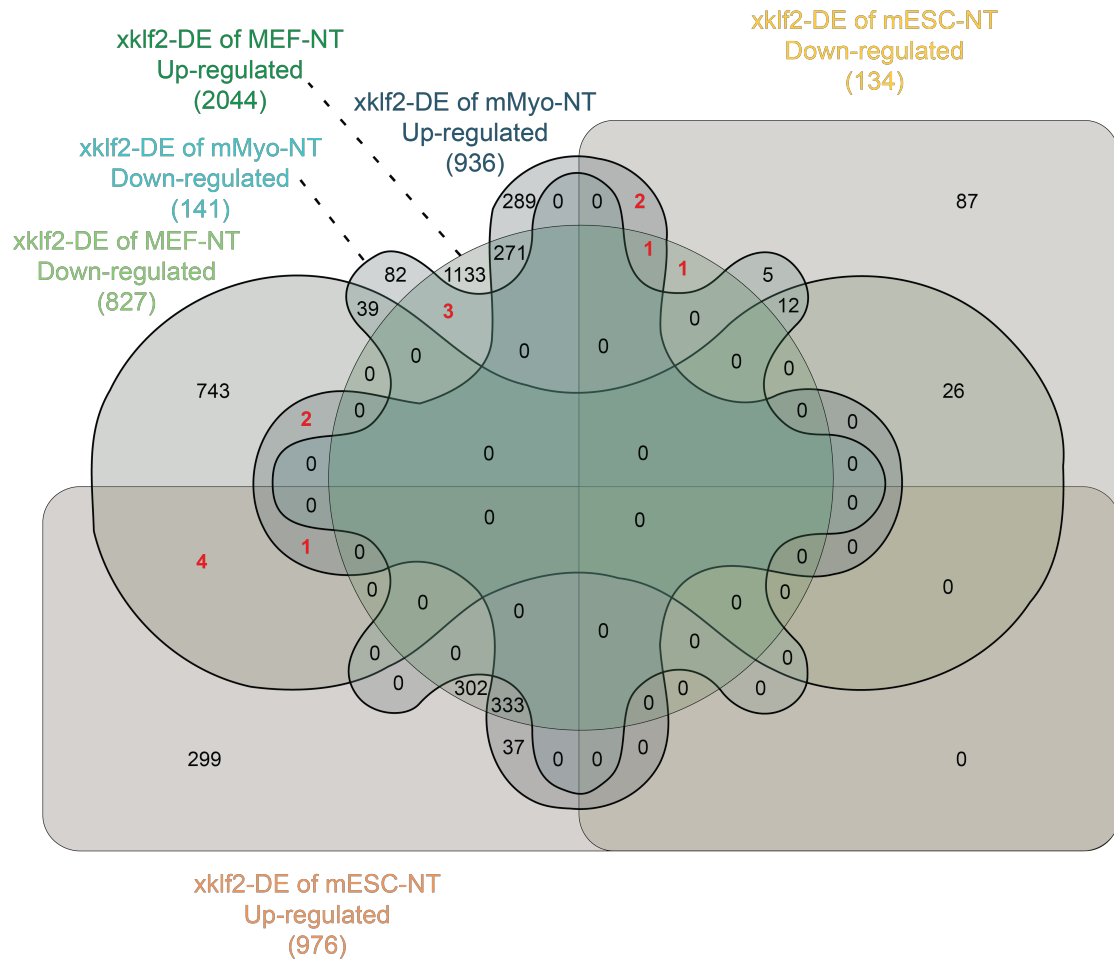


Figure 6.4.2 Regulation of xklf2-DE genes by xklf2-HA overexpression is one direction during Oocyte-NT. xklf2-DE genes are either up-regulated by xklf2-HA overexpression in all 3 cell types or are down-regulated by xklf2-HA overexpression in all 3 cell types. 14 xklf2-DE genes (in red) are up-regulated and down-regulated by xklf2-HA overexpression among 3 cell types.

6.4.2 xklf2-HA overexpression up-regulates expression of xklf2-DE genes to xklf2-oocyte level

To look further into up-regulation of xklf2-DE genes by xklf2-HA overexpression in different cell nuclei at Day 2 after Oocyte-NT, expression level of xklf2-DE genes in mESC-NT, MEF-NT and mMyo-NT are compared and \log_2FC of xklf2-DE genes between Control groups and xklf2-HA groups in mESC-NT, MEF-NT and mMyo-NT are compared (Figure 6.4.3, 6.4.4 and 6.4.5). For the results of genes down-regulated by xklf2-HA overexpression, it is described in Appendix X (page 365).

There are 333 xklf2-DE genes of mESC-NT, MEF-NT and mMyo-NT up-regulated by xklf2-HA overexpression at Day 2 after oocyte-NT (Figure 6.4.1.B). The means of expression level (FPKM) of these 333 xklf2-DE genes in mESC-NT, MEF-NT and mMyo-NT are 1.27, 1.63 and 0.12, respectively (Figure 6.4.3.A). The corresponding means of \log_2FC (xklf2-HA/Control) of these 333 xklf2-DE genes in mESC-NT, MEF-NT and mMyo-NT are 4.23, 3.53 and 5.06, respectively (Figure 6.4.3.B). Therefore, xklf2-HA overexpression up-regulates genes at lower expression level (FPKM) with higher \log_2FC significantly (Mean of FPKM: 1.63 of MEF-NT > 1.24 of mESC-NT > 0.12 of mMyo-NT; Mean of \log_2FC : 3.53 of MEF-NT < 4.23 of mESC-NT < 5.06 of mMyo-NT, Figure 6.4.3).

For xklf2-DE genes up-regulated in 2 cell types, it shows the same up-regulation of genes by xklf2-HA overexpression (Figure 6.4.4) as xklf2-DE genes up-regulated in all 3 cell types (Figure 6.4.3). While the expression

level (FPKM) is lower in mESC-NT, MEF-NT or mMyo-NT of Control groups, xklf2-HA overexpression up-regulates expression of these genes to fixed xklf2-oocyte levels with higher \log_2FC . Take up-regulated xklf2-DE genes of mESC-NT and MEF-NT as an example, while mean of expression level (FPKM) of xklf2-DE genes is 0.35 in mESC-NT and 1.16 in MEF-NT (Figure 6.4.4.A), mean of \log_2FC is 4.64 in mESC-NT and 3.17 in MEF-NT (Figure 6.4.4.B).

For xklf2-DE genes up-regulated in only 1 cell type (Figure 6.4.5), up-regulation of genes by xklf2-HA overexpression is the same as genes up-regulated in 2 and 3 cell types while the lower the expression level of genes in Control groups, the higher of \log_2FC is (Figure 6.4.3 and 6.4.4). It indicates each gene in transplanted cell nuclei is up-regulated by maternal factors and xklf2-HA overexpression to a fixed expression level, which is called xklf2-oocyte level and decided by the effects of maternal factors and xklf2-HA overexpression.

Notably, number of genes for \log_2FC comparison is sometimes different because some genes are resistant to be activated by xklf2-HA overexpression and are excluded from dataset. Therefore, although mean of expression level (FPKM) of up-regulated xklf2-DE genes of mESC-NT and MEF-NT is 1.16 in MEF-NT > 0.68 in mMyo-NT > 0.35 in mESC-NT (Figure 6.4.4.A), mean of \log_2FC is 2.51 in mMyo-NT < 3.17 in MEF-NT < 4.64 in mESC-NT (Figure 6.4.4.B).

Figure 6.4.3

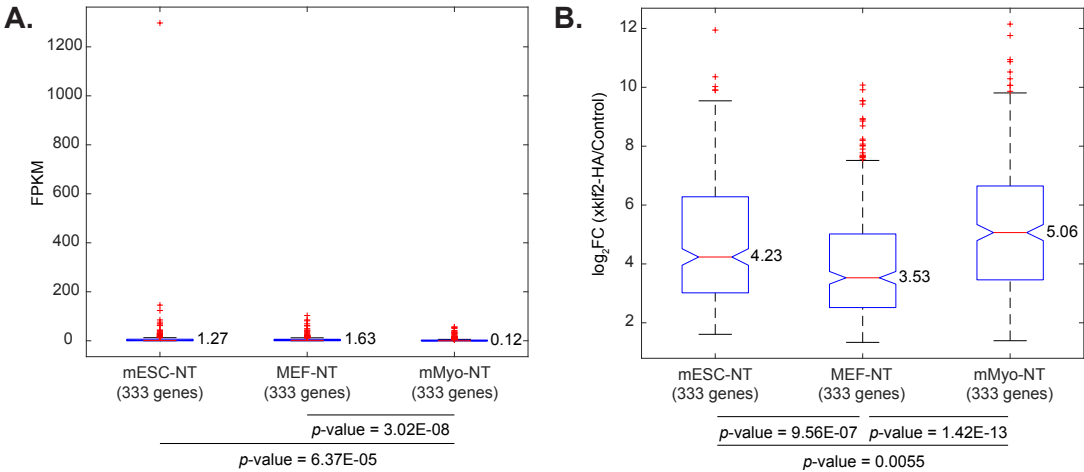


Figure 6.4.3 Expression level (FPKM) in the absence of xklf2-HA and $\log_2\text{FC}$ of shared up-regulated xklf2-DE genes in all 3 cell types.

Figure 6.4.4

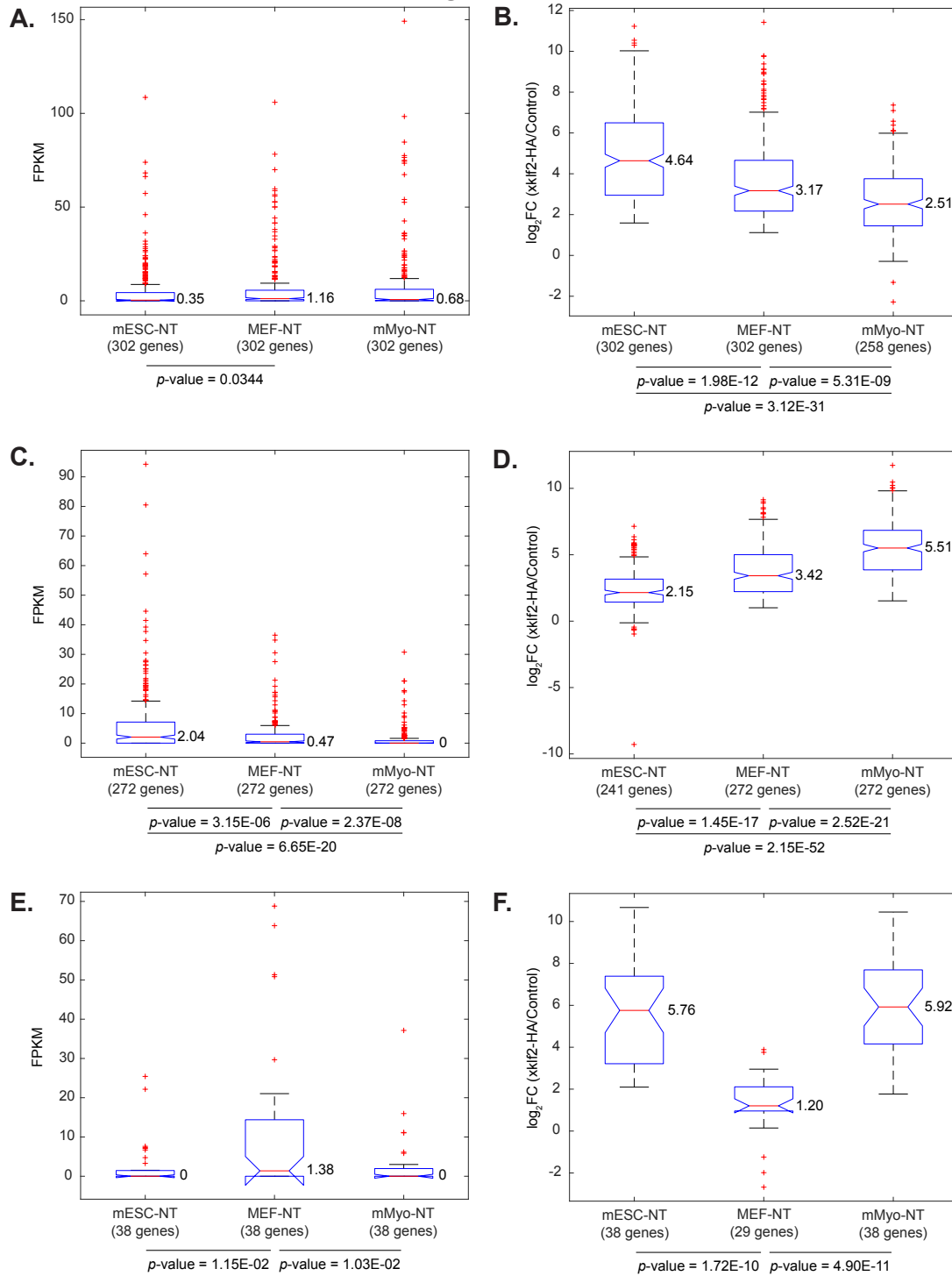


Figure 6.4.4 Shared up-regulated xklf2-DE genes in 2 cell types have lower expression level (FPKM) in the absence of xklf2-HA and up-regulated by xklf2-HA overexpression with higher \log_2FC than the 3rd cell type at Day 2 after Oocyte-NT.

(A, B) Up-regulated xklf2-DE genes shared between mESC-NT and MEF-NT

(C, D) Up-regulated xklf2-DE genes shared between MEF-NT and mMyo-NT

(E, F) Up-regulated xklf2-DE genes shared between mESC-NT and mMyo-NT

Figure 6.4.5

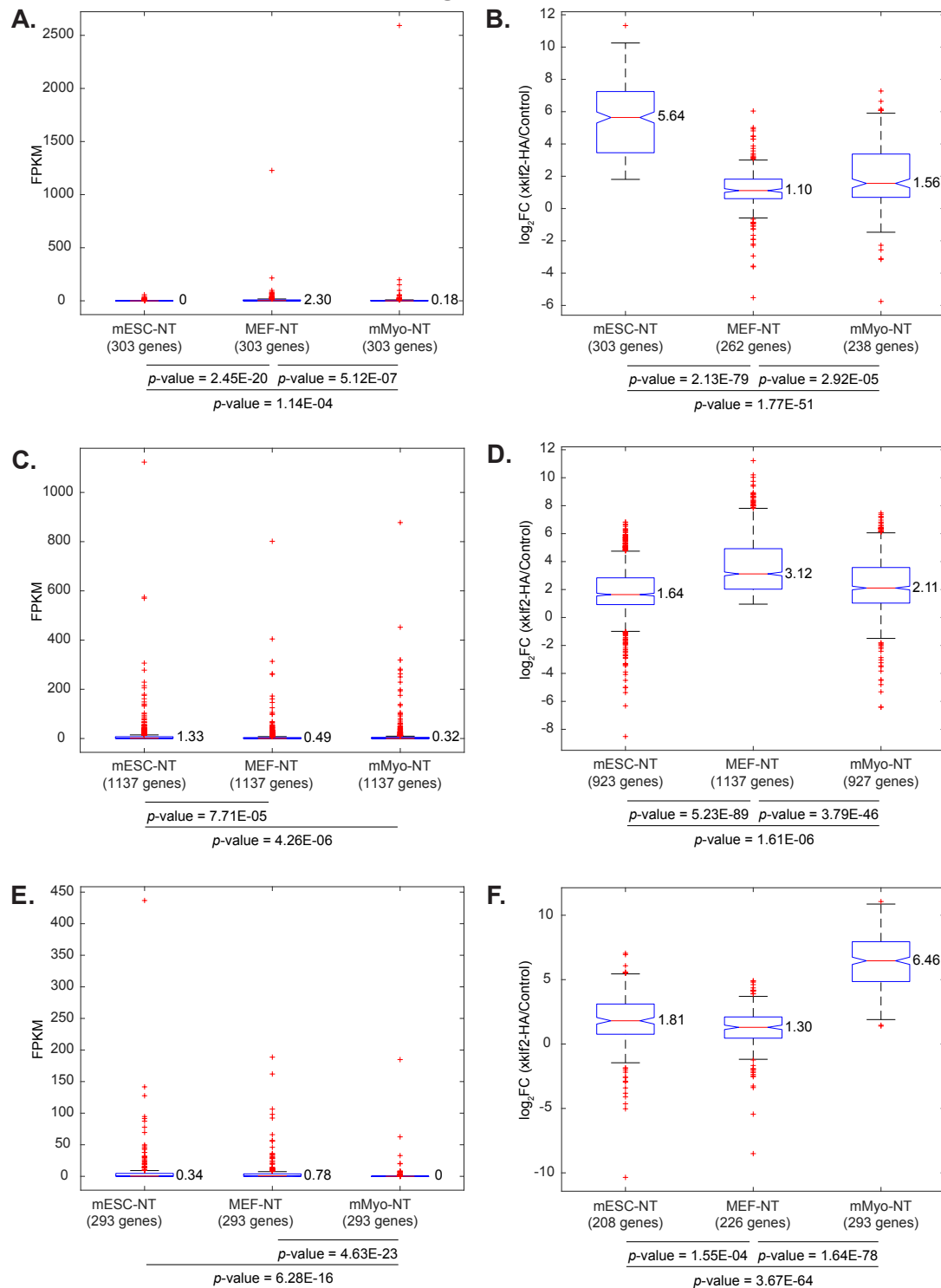


Figure 6.4.5 Up-regulated *xklf2*-DE genes in only one cell type have lower expression level in the absence of *xklf2*-HA and up-regulated by *xklf2*-HA overexpression with higher \log_2FC than the other 2 cell types at Day 2 after Oocyte-NT.

(A, B) Up-regulated *xklf2*-DE genes in only mESC-NT

(C, D) Up-regulated *xklf2*-DE genes in only MEF-NT

(E, F) Up-regulated *xklf2*-DE genes in only mMyo-NT

6.4.3 Summary

In *Xenopus* oocytes, gene regulation by exogenous TFs in transplanted cell nuclei is comprehensive because maternal factors also participate in gene regulation in transplanted cell nuclei during SCNR by oocytes¹⁴⁵.

In this section, it shows that *xklf2*-HA overexpression regulates different sets of *xklf2*-DE genes in mESC-NT, MEF-NT and mMyo-NT and *xklf2*-DE genes that are either up-regulated or down-regulated by *xklf2*-HA overexpression at Day 2 after Oocyte-NT. Maternal factors and *xklf2*-HA overexpression combinatorially determine which genes are regulated in transplanted cell nuclei. While expression level of *xklf2*-DE genes of transplanted cell nuclei is distant from *xklf2*-oocyte level, log2FC is higher when genes are up-regulated by *xklf2*-HA overexpression or log2FC is lower when genes are down-regulated by *xklf2*-HA overexpression.

All in all, different sets of *xklf2*-DE genes in mESC-NT, MEF-NT and mMyo-NT are decided by different expression levels of *xklf2* down stream genes as well as resistance of downstream genes against SCNR by oocytes or *xklf2*-HA overexpression. Therefore, all the *xklf2*-DE genes in mESC-NT, MEF-NT and mMyo-NT can be regulated by *xklf2*-HA overexpression by either up-regulating or down-regulating them to *xklf2*-oocyte level during SCNR by oocytes.

6.5 Up-regulated xklf2-DE genes by xklf2-HA overexpression during SCNR by oocytes are responsible for developmental processes and signaling pathways regulating pluripotency of stem cells

During induced pluripotency, it has been shown that differentiated cells are reprogrammed to be epigenetically and transcriptionally identical to ESCs by overexpression of Yamanaka factors, OSKM³⁵. In Chapter 5, it has been shown that transcriptomes of mESCs, MEFs and mMyos are reprogrammed to be almost identical at Day 2 after Oocyte-NT and chromatin structures of mESCs, MEFs and mMyos are forcibly changed by maternal factors of oocytes. Therefore, identical transcriptomes of reprogrammed cells are relevant to successful SCNR.

In Section 6.3 and 6.4, xklf2-HA overexpression has been shown to regulate expression of genes to xklf2-oocyte level in mESC-NT, MEF-NT and mMyo-NT during SCNR by oocytes. In this section, I would like to ask what functional interpretations are enriched for downstream genes of xklf2-HA during SCNR by oocytes and provide evidence for the role of xklf2 if it is one of maternal factors regulating genes during SCNR by oocytes/eggs?

6.5.1 The majority of xklf2-DE genes have mKlf2 binding motifs in their promoters and xklf2-DE genes are signaling pathways regulating pluripotency of stem cells are enriched for up-regulated xklf2-DE genes

To evaluate the way that xklf2-DE genes are regulated by xklf2-HA overexpression at Day 2 after Oocyte-NT, xklf2-DE genes are enriched for prediction of transcription factor binding sites in promoters (TSS \pm 3kb) via TRANSFAC database (Figure 6.5.1).

2821 in 3672 xklf2-DE genes regulated by xklf2-HA overexpression at Day 2 after Oocyte-NT are input into TRANSFAC database for enrichment (Figure 6.5.1.A). It shows that 2184 in 2821 xklf2-DE genes are enriched for mKlf2-binding motif: GGGGTGGKSN (p -value=5.47E-29, Figure 6.5.1.A). 1496 in 2821 xklf2-DE genes are enriched for mKlf2-binding motif: NGGGCGG (p -value=2.25E-24, Figure 6.5.1.A). 1693 in 2821 xklf2-DE genes are enriched for mKlf2-binding motif: CNCCACCCS (p -value=2.2E-34, Figure 6.5.1.A).

Among 2821 xklf2-DE genes, 2024 xklf2-DE genes are up-regulated and 797 xklf2-DE genes are down-regulated by xklf2-HA overexpression (Figure 6.5.1.B). By overlapping with xklf2-DE genes with mKlf2 binding sites, 90% of up-regulated xklf2-DE genes and 77% of down-regulated xklf2-DE genes have mKlf2 binding motifs in their promoters (Figure 6.5.1.B). For the rest of up-regulated and down-regulated xklf2-DE genes, they are probably regulated either by xklf2-HA binding on enhancers by xklf2-HA or by downstream genes of xklf2-HA.

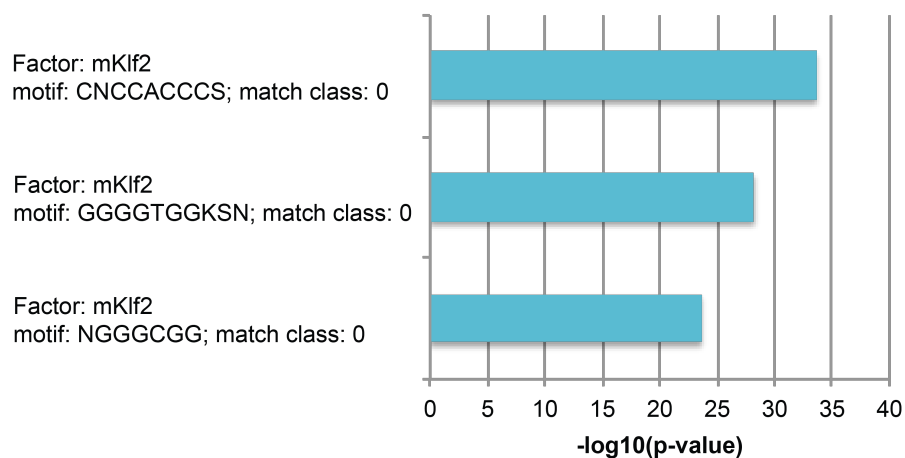
Notably, there is higher percentage of down-regulated *xklf2*-DE genes without *mKlf2* binding sites in their promoters than percentage of up-regulated *xklf2*-DE genes (Figure 6.5.1.B). It supports the hypothesis of section 6.2 that downstream genes of *xklf2*-HA are also involved in regulation of *xklf2*-DE genes because more genes are down-regulated in the presence of *xklf2*-HA overexpression from Day 1 to Day 2 after Oocyte-NT (Figure 6.2.2.C and Table 6.2.2, page 200 and 201) while *xklf2*-HA majorly plays as a transcriptional activator within one day after Oocyte-NT (Figure 6.2.1.C and Table 6.2.1, page 196 and 197).

The functional enrichment of up-regulated and down-regulated *xklf2*-DE genes was analysed by Gene ontology and KEGG pathway analysis in Appendix XI (page 369). Importantly, signaling pathways regulating pluripotency of stem cells (139 annotated genes in database) are enriched most significantly by up-regulated *xklf2*-DE genes and 37 up-regulated *xklf2*-DE genes are annotated in these signaling pathways ($p=6.23E-07$, Figure 10.11.4, page 380; these 37 genes are used later in Section 6.6).

Figure 6.5.1

A.

TRANSFAC
xklf2-DE genes of all 3 cell types



B.

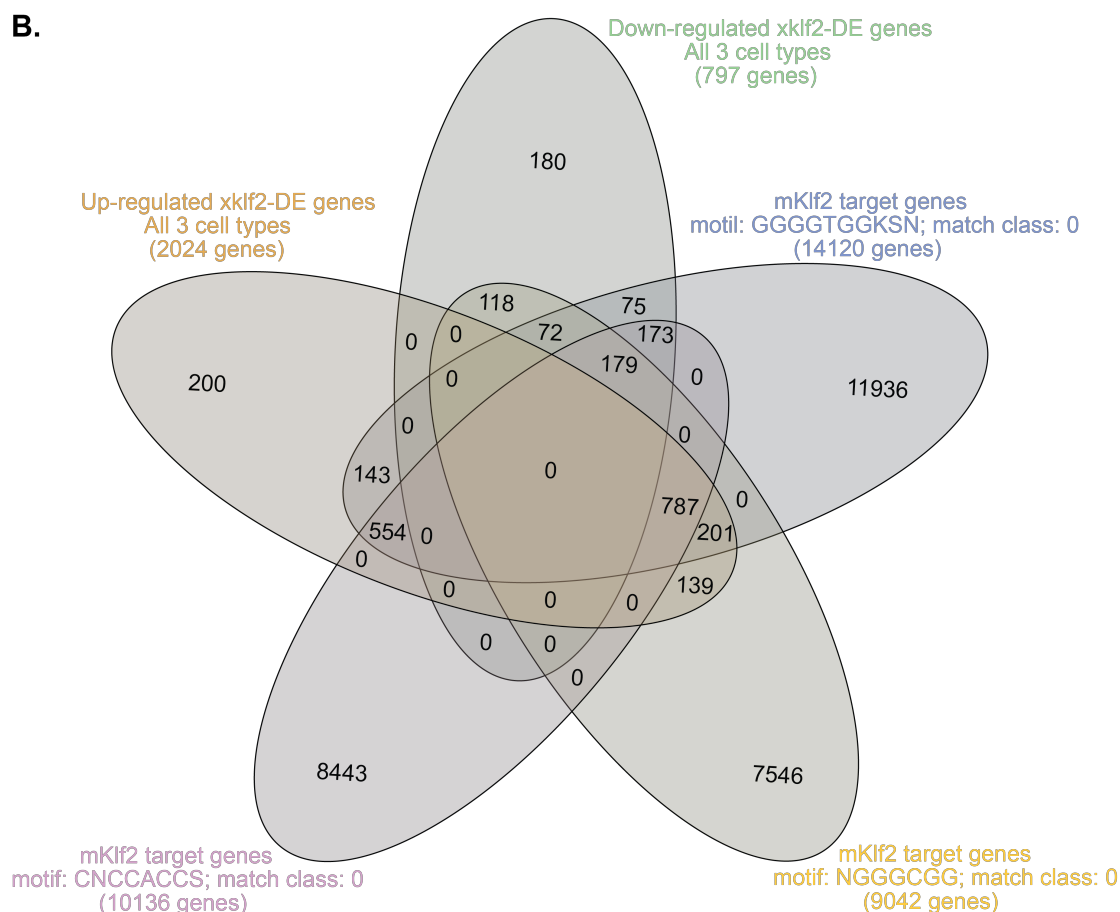


Figure 6.5.1 mKlf2 binding motif in promoters is enriched for all xklf2-DE genes of all 3 cell types regulated by xklf2-HA overexpression at Day 2 after Oocyte-NT. 90% of up-regulated xklf2-DE genes and 77% of down-regulated xklf2-DE genes of all 3 cell types have mKlf2 binding sites in their promoters.

6.5.2 Summary

In section 6.4, it has been shown that gene regulation by xklf2-HA overexpression in *Xenopus* oocytes depends on the donor cell types and xklf2-HA overexpression selectively regulate genes especially at nil or low level in reprogrammed transcriptomes of mESCs, MEFs and mMyos at Day 2 after Oocyte-NT. The selective regulation by xklf2-HA overexpression is caused by the fact that both maternal factors and xklf2-HA overexpression are involved in regulating expression of genes and xklf2-HA overexpression regulates those genes whose expression have not reached xklf2-oocyte level. Therefore, downstream genes of xklf2-HA during SCNR by oocytes should include all the xklf2-DE genes in mESCs, MEFs and mMyos.

In order to predict the functions of maternal xklf2, all the xklf2-DE genes in mESCs, MEFs and mMyos are collectively annotated for functional interpretation via TRANSFAC database, GO enrichment analysis and KEGG pathway enrichment analysis.

When analyzing xklf2-DE genes via TRANSFAC database, mKlf2 binding sites are significantly enriched in promoters (TSS \pm 3kb) of xklf2-DE genes. 90% of up-regulated xklf2-DE genes and 77% of down-regulated xklf2-DE genes have mKlf2 binding sites in their promoters. The rest 10% of up-regulated xklf2-DE genes and 23% of down-regulated xklf2-DE genes can reasonably be judged to be regulated by downstream genes of xklf2-HA.

Additionally, functions enriched for up-regulated xklf2-DE genes and promoted by xklf2-HA overexpression include more than 100 developmental processes, transcriptional activator and repressor activity via core promoter proximal region binding, analyzed by GO enrichment analysis. Importantly, KEGG pathway enrichment analysis shows that 39 up-regulated xklf2-DE genes are annotated to be part of signaling pathways regulating pluripotency of stem cells.

For down-regulated xklf2-DE genes, DNA, RNA and protein processing related functions are enriched via GO enrichment analysis. Furthermore, pathways for basic cellular functions are enriched for down-regulated xklf2-DE genes via KEGG pathway enrichment analysis, such as DNA replication, ribosome, nucleotide excision repair.

All in all, more than 87% of xklf2-DE genes have mKlf2 binding sites in their promoters and can be regulated by direct binding of xklf2-HA while binding sites of transcription factor homologs are often identical or very similar. If xklf2 is one of maternal factors in oocytes/eggs, it can activate genes, which are involved in developmental process, or activate genes, which are transcriptional activators or repressors during SCNR by oocytes.

6.6 xklf2-HA overexpression activates SCNR resistant genes in mESCs, MEFs and mMyos

In Chapter 5, the reprogrammable genes and cell-type specific genes during SCNR by oocytes are defined by expression level (FPKM) of genes in reprogrammed transcriptomes of mESCs, MEFs and mMyos. Reprogrammable genes are genes expressed at same level in all three reprogrammed transcriptomes of mESCs, MEFs and mMyos. Cell-type specific genes are genes, which are not expressed (FPKM=0) in one or two reprogrammed transcriptomes of mESCs, MEFs or mMyos. Among cell-type specific genes, while these genes are not expressed in all 3 cell types, expression of some cell-type specific genes are maintained by transcriptional machineries of transplanted cells and some cell-type specific genes are resistant to SCNR by oocytes due to inaccessible chromatin structure of transplanted cell types.

In Section 6.4, it has been shown that xklf2-HA overexpression regulates different sets of xklf2-DE genes in reprogrammed transcriptomes of mESCs, MEFs and mMyos because both maternal factors and xklf2-HA regulate expression of genes to xklf2-oocyte level. Here, I would like to ask if xklf2-HA overexpression could activate SCNR resistant genes in mESCs, MEFs and mMyos?

6.6.1 42~46% of xklf2-DE genes are reprogrammable genes and are regulated by both maternal factors and xklf2-HA overexpression

Cell-type specific genes include some genes, whose expression is maintained by transcriptional machineries of transplanted cells, and SCNR resistant genes, which can be regulated in one or two reprogrammed transcriptomes by maternal factors but resist to be activated by maternal factors in other two or one reprogrammed transcriptomes at Day 2 after Oocyte-NT. To evaluate the way that xklf2-HA overexpression regulates genes during SCNR by oocytes, reprogrammed transcriptomes of mESCs, MEFs and mMyos (FPKM>0) are compared with xklf2-DE genes in reprogrammed transcriptomes of mESC, MEFs and mMyos (mESC-NT, MEF-NT and mMyo-NT, Figure 6.6.1, 6.6.2 and 6.6.3).

Among xklf2-DE genes, 63% of xklf2-DE genes of mESC-NT are expressed in mESC-NT (FPKM>0) (703/1110, Figure 6.6.1.A), 77% of xklf2-DE genes of MEF-NT are expressed in MEF-NT (FPKM>0) (2202/2871, Figure 6.6.1.B) and 52% of xklf2-DE genes of mMyo-NT are expressed in mMyo-NT (FPKM) (562/1077, Figure 6.6.1.C). The rest 37%, 23% and 48% of xklf2-DE genes of mESC-NT, MEF-NT and mMyo-NT include 276, 359 and 362 cell-type specific genes, which are not expressed in mESC-NT, MEF-NT and mMyo-NT, and 131, 310 and 153 genes, which are not expressed in all 3 cell types, respectively (Figure 6.6.1). Therefore, more than 52% of xklf2-DE genes are already expressed in reprogrammed transcriptomes of mESCs, MEFs and mMyos.

For 11488 reprogrammable genes (FPKM>0, all 3 cell types), 505 xklf2-DE genes in mESC-NT, 1542 xklf2-DE genes in MEF-NT and 454 genes in mMyo-NT are reprogrammable genes and regulated by xklf2-HA overexpression at Day 2 after Oocyte-NT, which account for 46%, 54% 42% of xklf2-DE genes expressed in mESC-NT, MEF-NT and mMyo-NT (FPKM>0) (505/1100, Figure 6.6.1.A; 1542/2871, Figure 6.6.1.B; 454/1077, Figure 6.6.1.C). Hence, more than 42% of xklf2-DE genes, which is expressed in mESC-NT, MEF-NT and mMyo-NT are regulated by both maternal factors and xklf2-HA overexpression.

6.6.2 5~11% of cell-type specific genes resist to SCNR by oocytes but are activated by xklf2-HA overexpression within 2 days after Oocyte-NT

For 5960 cell-type specific genes, which are not expressed in mESC-NT but expressed in MEF-NT, mMyo-NT or both, 276 xklf2-DE genes in mESC-NT are activated by xklf2-HA overexpression and accounts for 5% of cell-type specific genes (Figure 6.6.1.A). For 3303 cell-type specific genes, which are not expressed in MEF-NT, 359 xklf2-DE genes in MEF-NT are activated by xklf2-HA overexpression and accounts for 11% of cell-types specific genes, which are not expressed in MEF-NT (Figure 6.6.1.B). For 6741 cell-type specific genes, which are not expressed in mMyo-NT, 362 xklf2-DE genes in MEF-NT are activated by xklf2-HA overexpression and accounts for 5% of cell-types specific genes, which are not expressed in mMyo-NT (Figure 6.6.1.C). Therefore, 5~11% of cell-type specific genes resist to be activated by maternal factors but are activated by xklf2-HA overexpression.

Interestingly, the third kind of genes is found when examining the effect of xklf2-HA overexpression. The third kind of genes is not expressed in all 3 cell types, resist to be activated by maternal factors and activated by xklf2-HA overexpression. For the third kind of genes, 131 genes are found in xklf2-DE genes of mESC-NT, 310 genes are found in xklf2-DE genes of MEF-NT and 153 genes are found in xklf2-DE genes of mMyo-NT (Figure 6.6.1).

6.6.3 Reprogrammed transcriptomes contain up-regulated and down-regulated xklf2-DE genes but only up-regulated xklf2-DE genes contribute to activate resistant genes

Within mESC-NT, MEF-NT and mMyo-NT (FPKM>0), 59%, 68% and 45% of up-regulated xklf2-DE genes are expressed in mESC-NT, MEF-NT and mMyo-NT and are up-regulated by xklf2-HA overexpression (571/976, mESC-NT, Figure 6.6.2.A; 1382/2044, MEF-NT, Figure 6.6.2.B; 422/936, Figure 6.6.2.C). Additionally, almost all down-regulated xklf2-DE genes are expressed in mESC-NT, MEF-NT and mMyo-NT without the addition of xklf2-HA (132/134, mESC-NT, Figure 6.6.3.A; 820/827, MEF-NT, Figure 6.6.3.B; 140/141, mMyo-NT, Figure 6.6.3.C). Therefore, xklf2-DE genes expressed in mESC-NT, MEF-NT and mMyo-NT are either up-regulated or down-regulated by xklf2-HA overexpression.

For genes that are not expressed in mESC-NT, MEF-NT and mMyo-NT, 405, 662 and 514 xklf2-DE genes are up-regulated and 2, 7 and 1 genes are down-regulated by xklf2-HA overexpression at Day 2 after Oocyte-NT (Figure

6.6.2 and 6.6.3), respectively. Hence, genes resist to be activated by maternal factors are activated by xklf2-HA overexpression.

Additionally, up-regulated xklf2-DE genes are activated by xklf2-HA overexpression to xklf2-oocyte level and the statistic analysis is in Appendix XII (page 385).

Figure 6.6.1

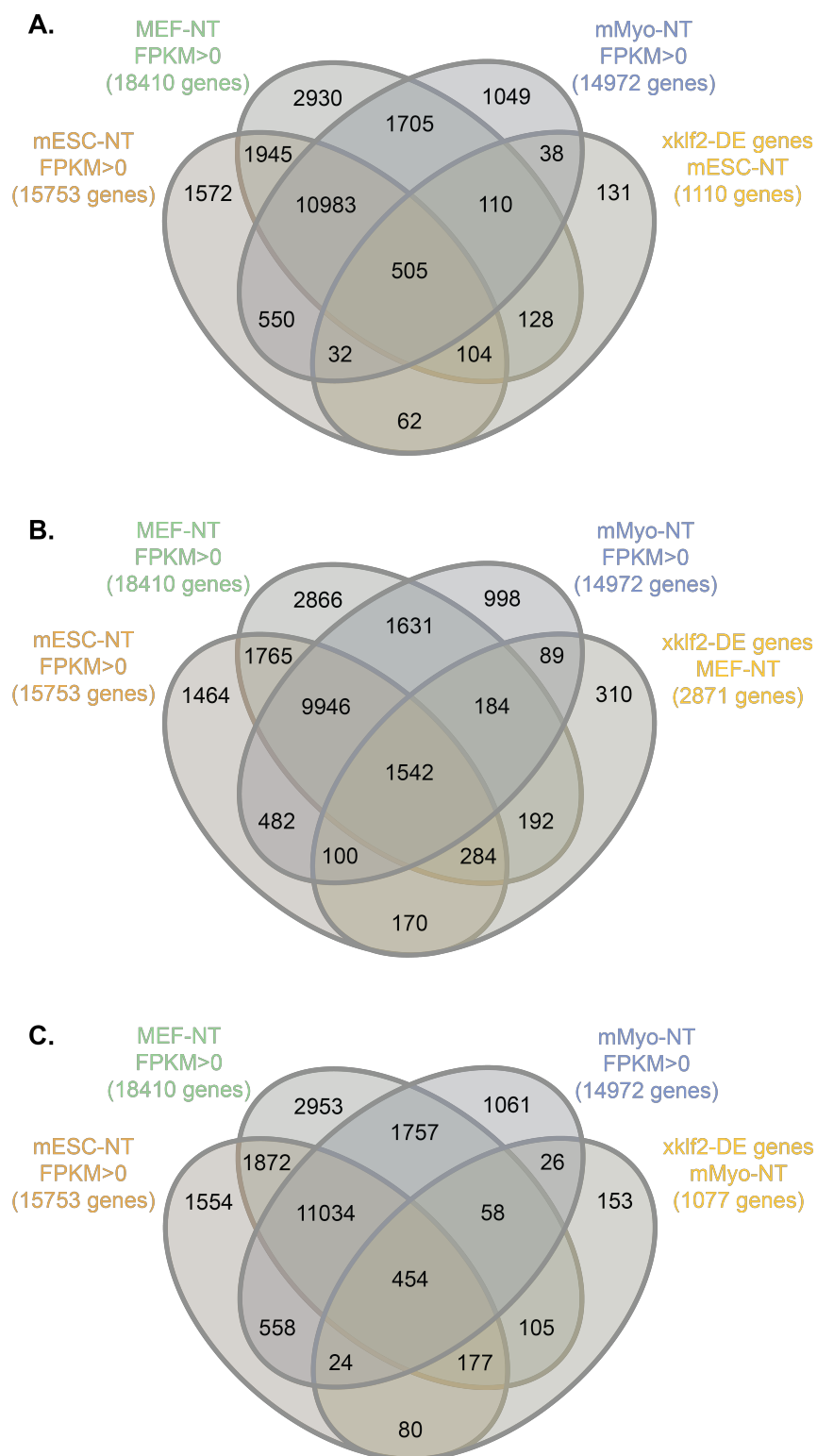


Figure 6.6.1 xklf2-HA overexpression regulates reprogrammable genes (expressed in all 3 cell types) and SCNR resistant genes, included cell-type specific genes (expressed in 1 or two cell types), at Day 2 after Oocyte-NT.

Figure 6.6.2

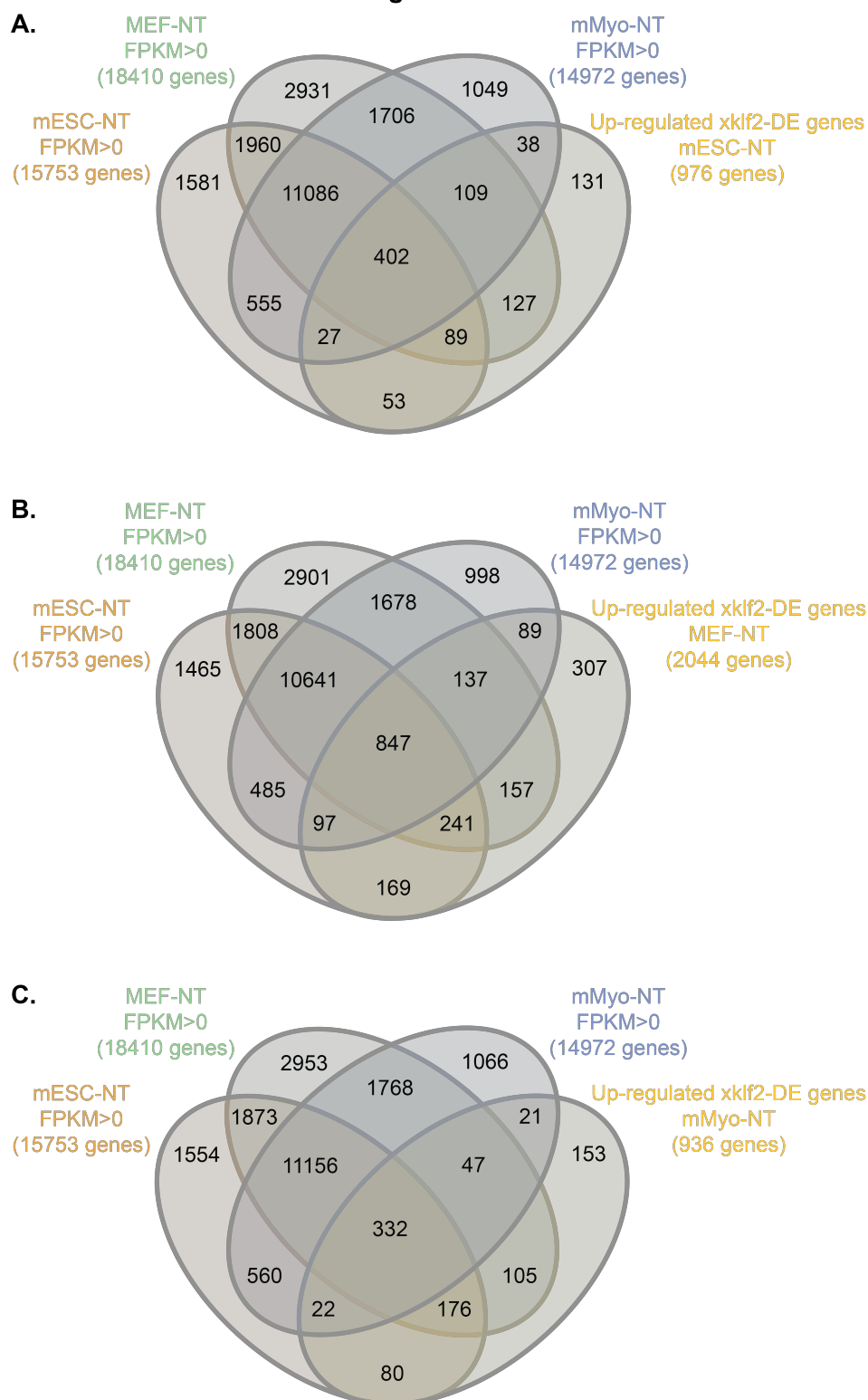


Figure 6.6.2 46, 54 and 42% of up-regulated xklf2-DE genes of mESC-NT (A), MEF-NT (B) and mMyo-NT (C) are reprogrammable genes, respectively. 28%, 17% and 39% of xklf2-DE genes of mESC-NT (A), MEF-NT (B) and mMyo-NT (C) are SCNR resistant genes of mESC-NT, MEF-NT and mMyo-NT, respectively. 13%, 15% and 16% of xklf2-DE genes of mESC-NT (A), MEF-NT (B) and mMyo-NT (C) are SCNR resistant genes of all 3 cell types.

Figure 6.6.3

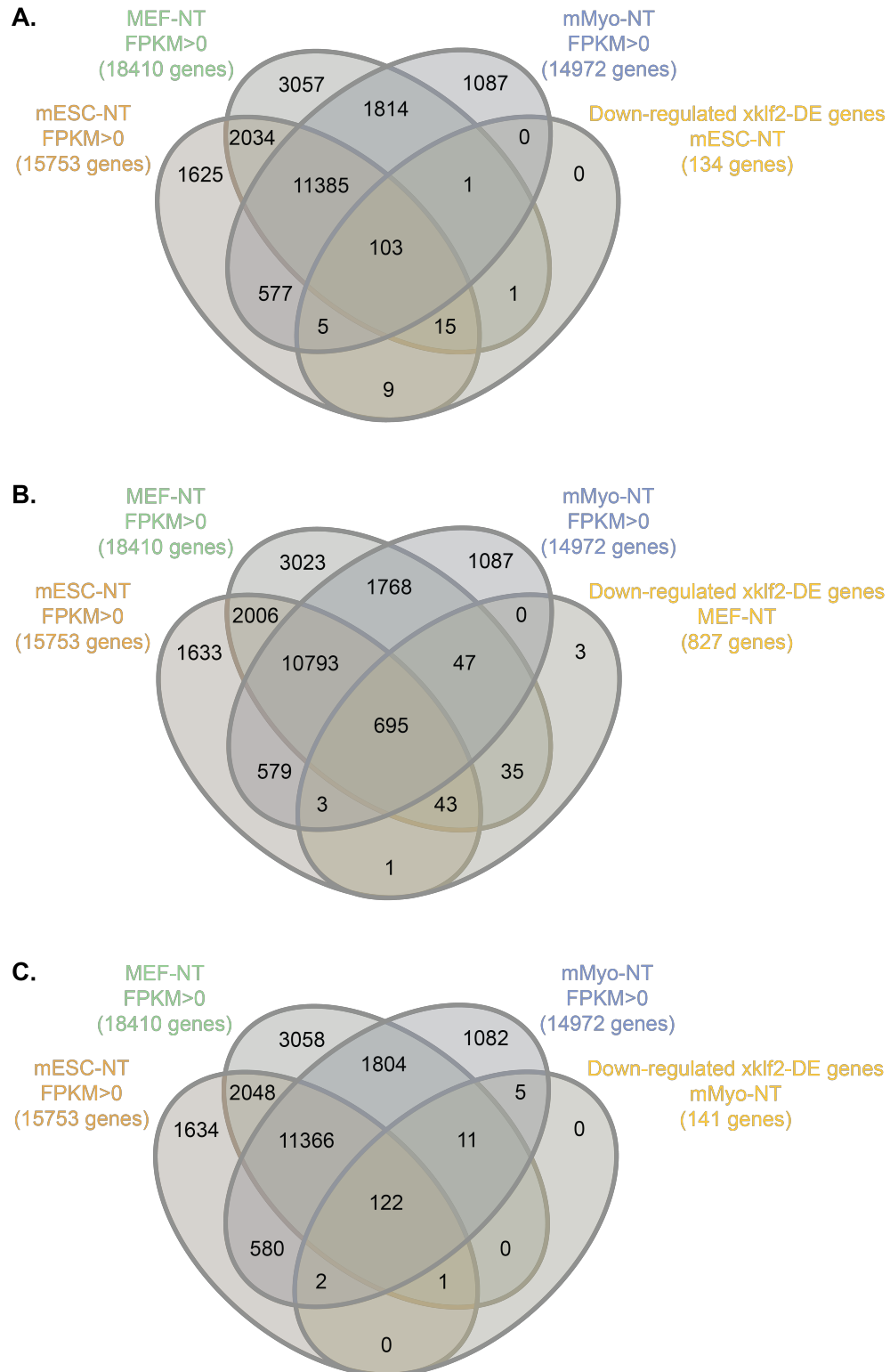


Figure 6.6.3 Down-regulated xklf2-DE genes are genes expressed already in mESC-NT (A), MEF-NT (B) and mMyo-NT (C) in the absence of xklf2-HA at Day 2 after Oocyte-NT.

6.6.4 xklf2-HA overexpression activates and regulates genes responsible for signaling pathways regulating pluripotency of stem cells

To evaluate the biological functions of up-regulation by xklf2-HA overexpression, the regulation of genes by xklf2-HA overexpression is applied in signaling pathways regulating pluripotency genes of stem cells. Previously, 39 up-regulated xklf2-DE genes are annotated to participate in signaling pathways regulating pluripotency genes of stem cells via KEGG pathway enrichment analysis and the regulation of them by xklf2-HA overexpression is shown in Table 6.6.1 and 6.6.2.

Since xklf2-HA overexpression can regulate both resistant genes and reprogrammable genes, the regulation of these two kinds of genes are separated in Table 6.6.1 and 6.6.2. In Table 6.6.1 and 6.6.2, resistant xklf2-DE genes are divided into three levels. Resistant xklf2-DE genes with $FPKM=0$, $0 < FPKM < 0.1$ and $0.1 < FPKM < 1$ are shown in orange, green and blue, respectively.

It has been shown that maternal factors, xklf2-HA and downstream genes of xklf2-HA are involved in gene regulation during SCNR by oocytes and the up-regulation of genes by xklf2-HA overexpression is stronger when the genes are at nil or low expression level. It shows clearly that the relationship between expression level of reprogrammed transcriptomes and up-regulation of genes by xklf2-HA overexpression in different cell types while the lower the FPKM is, the higher the log2FC is (Table 6.6.1 and 6.6.2).

Take Pou5f1 (Oct4) as an example. When FPKM of Pou5f1 is 31.9, 0 and 0 in mESC-NT, MEF-NT and mMyo-NT, \log_2FC of Pou5f1 is 2, 3 and 6 in the presence of overexpressed xklf2-HA (Table 6.6.1). In some cases, \log_2FC of resistant xklf2-DE genes are not shown in some cell types but these resistant xklf2-DE genes can be activated/regulated by xklf2-HA overexpression in other cell types. Because genes would be excluded for DE analysis when they do not pass 50% filtering, it may indicate regulatory elements of them reside heterochromatin where it is R-form chromatin for being totally inaccessible by TFs¹⁴⁹ (Table 6.6.1).

Therefore, xklf2-HA overexpression can promote signaling pathways regulating pluripotency of stem cells by up-regulating 39 genes responsible for these pathways, of which 16 genes are resistant to SCNR by oocytes.

Table 6.6.1

Expression level of SCNR resistant genes of *xklf2*-DE genes in the absense of *xklf2*-HA

A.

Gene	FPKM		
	mESC-NT	MEF-NT	mMyo-NT
Fgf2	0.0	0.0	0.0
Neurog1	0.0	0.8	0.0
Inhbc	0.0	1.4	0.0
Inhba	0.0	1.9	0.6
Wnt7b	0.0	2.0	9.0
Onecut1	0.1	0.0	0.0
Lefty2	0.1	0.5	0.0
Hoxb1	0.2	0.0	0.0
Wnt16	0.4	0.0	0.0
Wnt8b	0.5	0.0	0.8
Lhx5	0.8	0.9	0.0
Inhbe	1.5	1.0	0.0
Isl1	3.7	0.0	0.0
Zic3	4.3	0.0	0.0
Dlx5	6.5	7.5	0.0
Pou5f1	31.9	0.0	0.0

Regulation of SCNR resistant gene by *xklf2*-HA overexpression

B.

Gene	log ₂ FC		
	mESC-NT	MEF-NT	mMyo-NT
Fgf2	6	7	ND
Neurog1	7	4	ND
Inhbc	8	3	ND
Inhba	6	7	8
Wnt7b	5	1	0
Onecut1	4	6	6
Lefty2	4	3	7
Hoxb1	5	6	5
Wnt16	4	3	5
Wnt8b	3	4	4
Lhx5	5	5	5
Inhbe	3	3	7
Isl1	1	5	7
Zic3	1	5	6
Dlx5	5	2	5
Pou5f1	2	3	6

C.

Gene	FDR		
	mESC-NT	MEF-NT	mMyo-NT
Fgf2	1.12E-01	7.57E-04	ND
Neurog1	7.52E-02	3.52E-02	ND
Inhbc	5.22E-02	6.01E-02	ND
Inhba	4.24E-09	1.24E-30	1.68E-10
Wnt7b	9.46E-02	7.79E-01	9.73E-01
Onecut1	5.29E-01	8.97E-03	3.35E-02
Lefty2	5.87E-02	1.25E-02	3.75E-03
Hoxb1	2.46E-01	7.51E-07	1.76E-01
Wnt16	1.48E-01	1.26E-02	2.13E-01
Wnt8b	4.28E-01	1.12E-02	5.85E-02
Lhx5	6.48E-03	1.15E-08	2.52E-02
Inhbe	2.36E-01	3.52E-03	4.63E-03
Isl1	6.77E-01	3.54E-02	4.01E-02
Zic3	8.35E-01	6.39E-03	1.18E-01
Dlx5	3.60E-06	1.28E-04	3.75E-03
Pou5f1	5.15E-02	3.65E-03	1.10E-02

Table 6.6.1 SCNR resistant genes of *xklf2*-DE genes (all 3 cell types) responsible for “signaling pathways regulating pluripotency of stem cells (KEGG pathway)” are activated strongly by *xklf2*-HA overexpression with high log₂FC. Expression level (FPKM) in the absense of *xklf2*-HA at Day 2 after Oocyte-NT is shown and FPKM=0 is in orange, 0<FPKM<0.1 is in green and 0.1<FPKM<1 is in blue. Log₂FC and FDR are shown for genes regulated by *xklf2*-HA overexpression at Day 2 after Oocyte-NT.

Table 6.6.2

Expression level of reprogrammable genes of *xklf2*-DE genes in the absence of *xklf2*-HA

A.

Gene	FPKM		
	mESC-NT	MEF-NT	mMyo-NT
Wnt10a	0.1	3.7	2.4
Pik3cd	0.3	1.0	0.3
Pik3r1	0.6	3.6	4.2
Igf1r	0.6	0.4	0.0
Wnt2b	3.0	6.2	2.6
Stat3	4.0	3.2	1.9
Dvl3	7.8	6.0	8.1
Smad1	7.9	2.3	4.1
Kat6a	9.6	19.4	6.4
Otx1	12.1	13.2	2.7
Wnt6	13.3	7.7	3.8
Fzd10	13.4	0.1	5.0
Id3	15.4	60.7	31.6
Id2	15.8	20.3	39.3
Jak3	17.8	1.8	0.4
Id1	18.2	10.7	13.0
Bmp4	22.2	44.2	12.5
Hand1	22.9	2.4	0.1
Zfhx3	30.5	1.8	7.0
Pax6	80.5	30.5	10.2
Sox2	1297.4	69.2	4.2

Regulation of reprogrammable genes by *xklf2*-HA overexpression

B.

Gene	log ₂ FC		
	mESC-NT	MEF-NT	mMyo-NT
Wnt10a	7	4	4
Pik3cd	1	2	1
Pik3r1	4	2	2
Igf1r	1	2	5
Wnt2b	1	1	3
Stat3	1	1	3
Dvl3	2	1	1
Smad1	2	4	4
Kat6a	2	3	2
Otx1	2	3	1
Wnt6	3	2	2
Fzd10	1	3	-1
Id3	2	1	2
Id2	2	2	1
Jak3	2	2	3
Id1	2	1	3
Bmp4	3	2	2
Hand1	4	3	5
Zfhx3	0	2	1
Pax6	1	1	2
Sox2	3	1	5

C.

Gene	FDR		
	mESC-NT	MEF-NT	mMyo-NT
Wnt10a	1.08E-09	5.02E-09	2.42E-01
Pik3cd	7.40E-01	9.87E-02	6.77E-01
Pik3r1	7.93E-04	8.90E-04	3.24E-02
Igf1r	7.83E-01	6.73E-02	1.49E-01
Wnt2b	3.39E-01	6.12E-02	5.02E-02
Stat3	7.63E-01	1.49E-01	7.00E-02
Dvl3	5.15E-02	3.62E-01	6.48E-01
Smad1	2.82E-02	1.49E-07	4.55E-05
Kat6a	2.93E-02	7.80E-07	1.16E-02
Otx1	3.94E-02	9.35E-07	6.94E-01
Wnt6	9.68E-03	2.25E-03	5.08E-01
Fzd10	7.96E-01	2.92E-03	7.41E-01
Id3	2.66E-02	5.18E-02	3.15E-02
Id2	2.30E-01	3.61E-03	2.91E-01
Jak3	1.41E-01	3.46E-02	6.23E-02
Id1	1.94E-01	4.73E-02	2.66E-02
Bmp4	3.49E-05	1.85E-02	1.36E-01
Hand1	1.07E-02	5.11E-07	1.21E-02
Zfhx3	8.97E-01	5.38E-02	6.43E-01
Pax6	2.04E-01	3.98E-02	1.99E-02
Sox2	5.02E-02	1.99E-02	6.37E-04

Table 6.6.2 Reprogrammable genes of *xklf2*-DE genes (all 3 cell types) responsible for “signaling pathways regulating pluripotency of stem cells (KEGG pathway)” are activated mildly by *xklf2*-HA overexpression with high log₂FC. Expression level (FPKM) in the absence of *xklf2*-HA at Day 2 after Oocyte-NT is shown and FPKM=0 is in orange, 0<FPKM<0.1 is in green and 0.1<FPKM<1 is in blue. Log₂FC and FDR are shown for genes regulated by *xklf2*-HA overexpression at Day 2 after Oocyte-NT.

6.6.5 Summary

In Chapter 5, it has been shown that there are two kinds of genes in reprogrammed transcriptomes of mESCs, MEFs and mMyos. One is reprogrammable genes, which are expressed in all 3 reprogrammed transcriptomes and regulated by maternal factors to the same level. The other kind of genes are cell-type specific genes and they are expressed in one or two cell types, rather than expressed in all 3 cell types. Cell-type specific genes contain genes, whose expression is maintained by transplanted cells, and SCNR resistant genes, which resist to be activated by maternal factors. In this section, it shows that *xklf2*-HA overexpression helps to recognize SCNR resistant genes by activating and regulating SCNR genes to *xklf2*-oocyte level.

By comparing reprogrammable genes, SCNR resistant genes and *xklf2*-DE genes, it shows that 42~46% of *xklf2*-DE genes are reprogrammable genes. Additionally, 5~11% of cell-type specific genes are recognized as SCNR resistant genes and are activated by *xklf2*-HA overexpression. In addition to that, approximately 300 *xklf2*-DE genes are shown to be silent in all 3 cell types but activated by *xklf2*-HA overexpression. Importantly, *xklf2*-HA overexpression can regulates reprogrammable genes and activate SCNR resistant genes.

Furthermore, activation of SCNR resistant genes is contributed by *xklf2*-HA up-regulation, rather than down-regulation. Maternal factors, *xklf2*-HA and downstream genes of *xklf2*-HA are involved in regulating expression of *xklf2*-DE genes to *xklf2*-oocyte level.

When applying up-regulation of genes by xklf2-HA overexpression in signaling pathways regulating pluripotency of stem cells, it shows that xklf2-HA overexpression can promote these signaling pathways by up-regulating 39 genes responsible for these pathways. Importantly, 16 genes of them resist to SCNR by oocytes and are activated by xklf2-HA overexpression.

All in all, xklf2-HA overexpression is shown to activate SCNR resistant genes in reprogrammed transcriptomes and regulate expression of xklf2-DE genes to xklf2-oocyte level cooperatively with maternal factors and downstream genes. Moreover, xklf2-HA overexpression activates and up-regulates 37 xklf2-DE genes, which participate in signaling pathways regulating pluripotency of stem cells.

6.7 Conclusions

In development, chromatin structures of cells are subjected to change to receive extracellular stimuli, which resulting new transcriptomes to be formed in the new cell types. During SCNR by oocytes, chromatin structures of transplanted cells are forcibly changed by maternal factors and nearly identical transcriptomes are formed at Day 2 after Oocyte-NT. However, SCNR resistant genes in transplanted cells are not always activated by maternal factors and it would lead to unsuccessful SCNR by oocytes. Since overexpression of transcription factors are also used to change cell identities, I would like to ask if overexpression of transcriptions can activate SCNR resistant genes and how transcription factors regulate genes during SCNR by oocytes?

Time-dependent effect of xklf2-HA overexpression has shown that most of xklf2-HA downstream genes are regulated by xklf2-HA overexpression within one day after Oocyte-NT. While more xklf2-HA downstream genes up-regulated by xklf2-HA overexpression than down-regulated by xklf2-HA overexpression within 1 day after Oocyte-NT but more xklf2-HA downstream genes are down-regulated than up-regulated by xklf2-HA overexpression from Day 1 to Day 2 after Oocyte-NT, it indicates downstream genes of xklf2-HA, especially repressors, are also involved in gene regulation by xklf2-HA overexpression.

When comparing effects of xklf2-HA overexpression on different cell types, xklf2-HA overexpression up-regulates expression of xklf2-DE genes, mostly at

nil or low expression level in reprogrammed transcriptomes of tested cell types, to *xklf2*-oocyte level at Day 2 after Oocyte-NT. When comparing up-regulated and down-regulated in all 3 cell types, it shows that the regulation of genes by *xklf2*-HA overexpression is one direction and *xklf2*-DE genes are either up-regulated or down-regulated by *xklf2*-HA overexpression. Additionally, mKlf2 binding motifs are enriched for *xklf2*-DE genes and shows regulation of *xklf2*-DE genes is highly possible through direct binding in their promoters.

Additionally, the functions of *xklf2* downstream genes can be annotated to the functions promoted by *xklf2*-HA overexpression during SCNR by oocytes. Through Gene ontology enrichment analysis, up-regulated *xklf2*-DE genes are annotated to developmental processes, transcriptional activators and repressors. Additionally, signaling pathways regulating pluripotency of stem cells are enriched from up-regulated *xklf2*-DE genes via KEGG pathway enrichment analysis and it indicates these pathways are promoted by *xklf2*-HA expression during SCNR by oocytes.

Finally, *xklf2*-HA overexpression can significantly regulate both reprogrammable genes and cell-type specific genes in reprogrammed transcriptomes of mESCs, MEFs and mMyos. Importantly, *xklf2*-HA overexpression helps to recognize SCNR resistant genes among cell-type specific genes and is shown to activate/regulate SCNR resistant genes to *xklf2*-oocyte level, including 16 SCNR resistant genes responsible for signaling pathways regulating pluripotency of stem cells.

Chapter 7 The effect of the pioneer transcription factor, mFoxa1-HA, on somatic cell nuclear reprogramming by *Xenopus* oocytes

7.1 Introduction

7.1.1 Background

In Chapters 5 and 6, I have shown that oocyte factors can reprogram transcriptomes of transplanted cells to an oocyte-steady state and that overexpression of *xklf2*-HA can activate target genes that are resistant to oocyte factors genome-wide. In this chapter, I ask if the pioneer transcription factor, mouse *Foxa1*-HA (mFoxa1-HA), can activate pluripotency genes and neurogenic genes through its pioneer transcription factor activity.

The difference between Yamanaka factors and mFoxa1 is that Yamanaka factors are important in early development since they are major factors to induce pluripotency whereas mFoxa1 is well known as a pioneer factor to access to and open the closed chromatin before the recruitment of other cofactors¹⁴⁹. Furthermore, it has been shown that *Xenopus foxa1* (*xfoxa1*) is activated after MBT in embryos^{140,141}. Therefore, it would be possible that there is no redundant factor like the role of *xfoxa1* or mFoxa1 in *Xenopus* oocytes and overexpression of mFoxa1-HA can induce expression of neurogenic genes, which are not induced by oocyte factors. However, it would be also possible that mFoxa1-HA, as a lineage-specific factor¹⁴⁹⁻¹⁵¹, may not

work as well as Yamanaka factors do due to the lack of cofactors and proper post-translational modifications.

7.1.2 Experimental design

In this chapter, the experiment procedures are similar to Chapter 3 and 4, except that I overexpressed mFoxa1-HA 24 hours before Oocyte-NT (Figure 7.1). In Section 7.2, I used a similar approach to overexpress mFoxa1-HA proteins and evaluate how it is produced in the *Xenopus* oocytes (Figure 7.1.A). As mentioned in Chapter 3, I overexpressed mFoxa1-HA in *Xenopus* oocytes and examined if mFoxa1-HA overexpression affects expression of pluripotency genes and neurogenic genes^{150,151} (Figure 7.1.B, Section 7.3 and 7.4). To compare the effect of oocyte factors and mFoxa1-HA overexpression on neurogenic genes, I did a time-course observation from 0 hour to 48 hours after Oocyte-NT (Figure 7.1.B, Section 7.4).

Figure 7.1

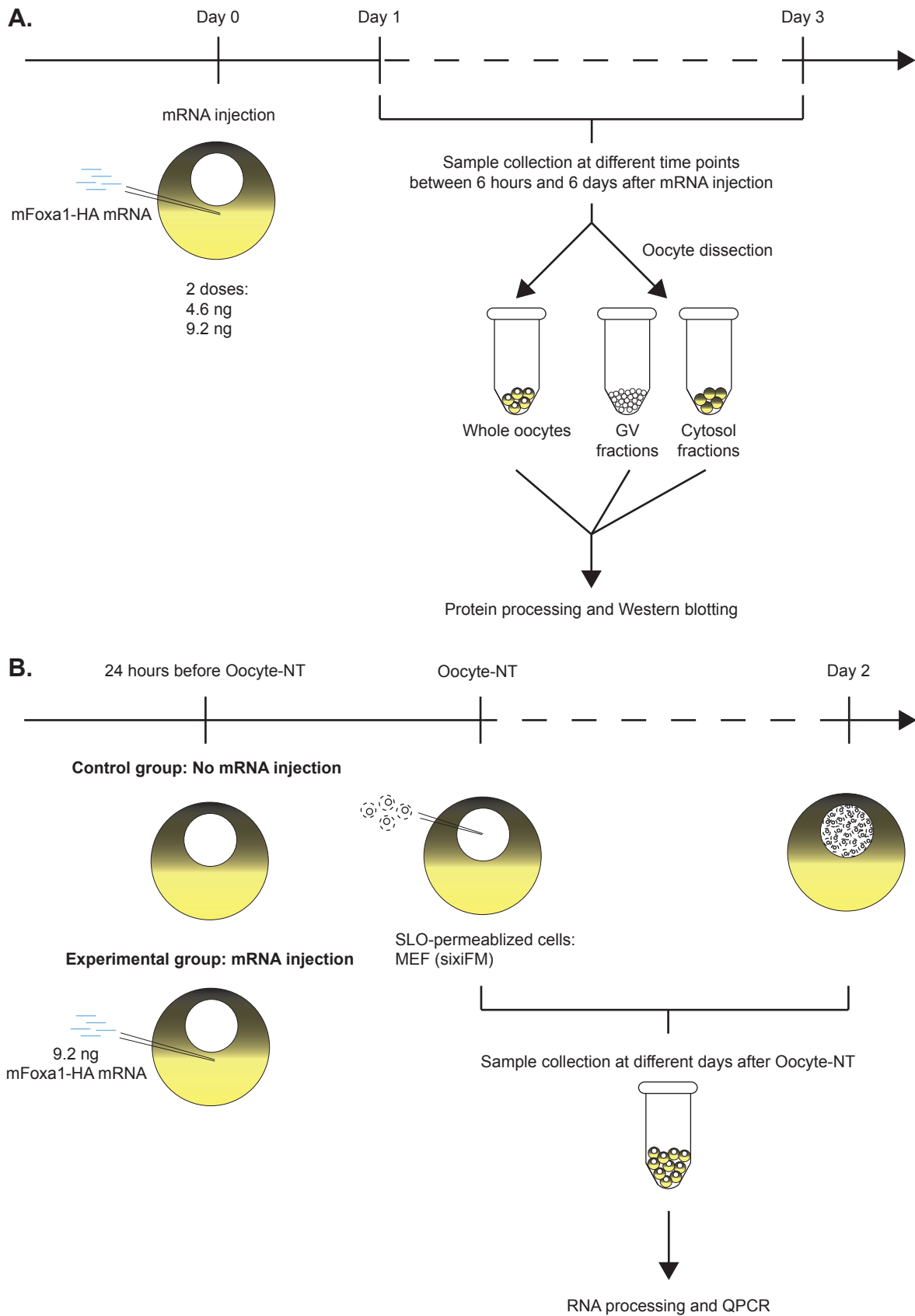


Figure 7.1 Sample preparations of Western blotting (A) and QPCR (B) for evaluating the production and effect of mFoxa1-HA overexpression are shown.

7.2 mFoxa1-HA accumulates in the GV dose-dependently and SUMOylation of mFoxa1-HA is observed in the *Xenopus* oocytes

To evaluate the production of mFoxa1-HA proteins, samples of different cellular compartments (whole oocytes, cytosol fractions and GV fractions) were collected at different days (Day 1, 2 and 3) after mFoxa1-HA mRNA injection (4.6 and 9.2 ng) and run on the same blot (Figure 7.2.1). The anti-HA antibody detects two obvious bands on the blot around the 55kD protein marker where the mFoxa1-HA should be (54kD = 49kD for mFoxa1 + 5kD for HA tag).

Figure 7.2.1

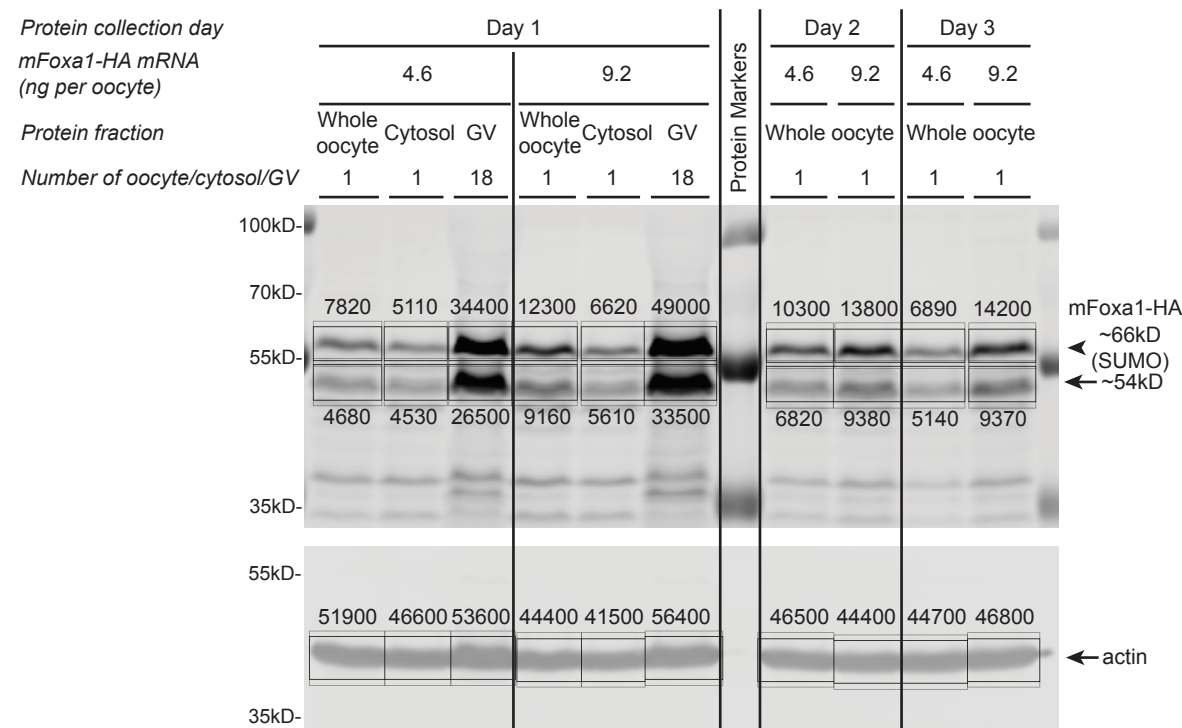


Figure 7.2.1 Signal intensity of mFoxa1-HA and actin at different conditions is shown.

The mFoxa1-HA proteins (upper black arrow and black arrowhead) are detected by anti-HA antibody and the actin proteins (lower black arrow) are detected by the anti-actin antibody. The signal intensity of mFoxa1-HA proteins and actin proteins are shown next to the quantification rectangles.

7.2.1 mFoxa1-HA proteins accumulate in the GV and increase time- and dose-dependently in different cellular compartments

To ask whether the mFoxa1-HA proteins will accumulate in the GVs, I added up the signals of 54kD bands and 66kD bands to make the bar chart of Figure 7.2.2.A. When injecting the oocytes with 4.6 ng and 9.2 ng of mFoxa1-HA mRNA and collecting the samples 1 day later, the mFoxa1-HA signals from 1 whole oocyte is 3.85 times ($=12500/3205$, 4.6 ng mFoxa1-HA mRNA) and 4.68 times ($=21460/4583$, 9.2 ng mFoxa1-HA mRNA) more than the signals from 1 GV fraction (Figure 7.2.2.A). Considering the volume ratio of 1 GV to 1 oocyte is 1/30 and 1 GV contains 1/4~1/5 of mFoxa1-HA proteins in whole oocytes, the mFoxa1-HA proteins must accumulate in the GVs. In addition to that, the more mFoxa1-HA proteins are made in the cytosol, the more mFoxa1-HA proteins are transported in the GVs.

Apart from these points, the actin signals in different fractions do not vary too much when the amount of mFoxa1-HA mRNA is doubled (Figure 7.2.1 and 7.2.2.B). When statistically analyzing the signal intensity of actin in different oocyte fractions and whole oocytes regardless of the amount of mFoxa1-HA mRNA and the sample collection days (Figure 7.2.2.C), the average signals intensity is 46450 for 1 whole oocyte (\pm SD=2874, 4.6 and 9.2 ng mRNA, Day 1, 2 and 3), 44050 for 1 cytosol fraction (\pm SD=3606, 4.6 and 9.2 ng mRNA, Day 1) and 3056 for 1 GV fraction (\pm SD=110, 4.6 and 9.2 ng mRNA, Day 1). The average signal intensity of 1 whole oocyte (46450) almost equals the average signal intensity of 1 cytosol fraction (44050) plus 1 GV fraction (3056). This means the detection of actin is very accurate within the signals from

3056 to 46450 and no dilution series of samples is needed for relative protein quantification here.

For an easier description of the fractionation difference, a stacked bar chart can be made (Figure 7.2.3.A and B) and the signal intensity can be divided by the whole oocyte sample and shown as percentages of relative signal intensity. It can help to examine what percent of protein signals in 1 whole oocyte is distributed to 1 cytosol fraction and 1 GV fraction. For example, in 4.6 ng mFoxa1-HA mRNA groups (left part, Figure 7.2.3.A), the mFoxa1-HA signals from 1 whole oocyte (100%) almost equals to the signals from 1 cytosol fraction (77%) plus the signals from 1 GV fraction (26%). 26% of the mFoxa1-HA signals in 1 whole oocyte is distributed to 1 GV fraction. When doubling the mFoxa1-HA mRNA to 9.2 ng per oocyte (right part, Figure 7.2.3.A), the signal percentages of 1 whole oocyte increase from 100% (4.6 ng group) to 172% (9.2 ng group) and the signal distribution to 1 GV is around 27% [= $37\% / (37\% + 98\%)$], which is almost the same as the distribution of 4.6 ng group (26%). Therefore, the distribution of mFoxa1-HA proteins to the GVs is not affected by the total amount of mFoxa1-HA proteins in the whole oocytes. Notably, in the 9.2 ng mFoxa1-HA mRNA groups, the signal percentages of 1 whole oocyte (172%) is ~1.3 times more than the signal percentages of 1 cytosol fraction (37%) plus 1 GV fraction (98%) and the uneven values would be acceptable due to the variable sizes of *Xenopus* oocytes, which could be minimized by collecting more oocytes to even the values further.

If checking the actin signals (Figure 7.2.3.B), the signals from 1 whole oocyte also roughly equal the signals from 1 cytosol fraction plus the signals from 1 GV fraction in both 4.6 ng mRNA groups and 9.2 ng mRNA groups. Additionally, because the actin signals from 1 GV fraction is $\sim 1/15$ of the 1 whole oocyte values (Figure 7.2.2.C and 7.2.3.B) and because the actin signals are not distributed to 1 GV fraction and 1 cytosol fraction with regard to the ratio of volume (GV:cytosol = $\sim 1/30$), it is not suitable to normalize the signals of proteins with signals of actin when comparing proteins in different fractions.

Then, I evaluated the dose- and time-dependency of mFoxa1-HA protein production and compared the mFoxa1-HA signals in whole oocyte samples at Day 1, 2 and 3 after mRNA injection (4.6 and 9.2 ng) (Figure 7.2.1 and 7.2.4). Firstly, the signals of two mFoxa1-HA bands are added up and normalized to the value of 4.6 ng group at Day 1 to make the line chart (Figure 7.2.4.A). At Day 1, the mFoxa1-HA signals of 9.2 ng mRNA group is 1.7 times more than the signals of 4.6 ng mRNA group, which is also validated previously (Figure 7.2.2.A and 7.2.3.A). At Day 2 and Day 3, the mFoxa1-HA signals of 9.2 ng group is 1.4 and 2 times more than the signals of 4.6 ng group, respectively (Figure 7.2.4.C). Since the amount of mFoxa1-HA proteins does not change greatly from Day 1 to Day 3 after mRNA injection without sign of protein degradation, this time period is suitable for performing Oocyte-NT (Figure 7.2.4.C).

Figure 7.2.2

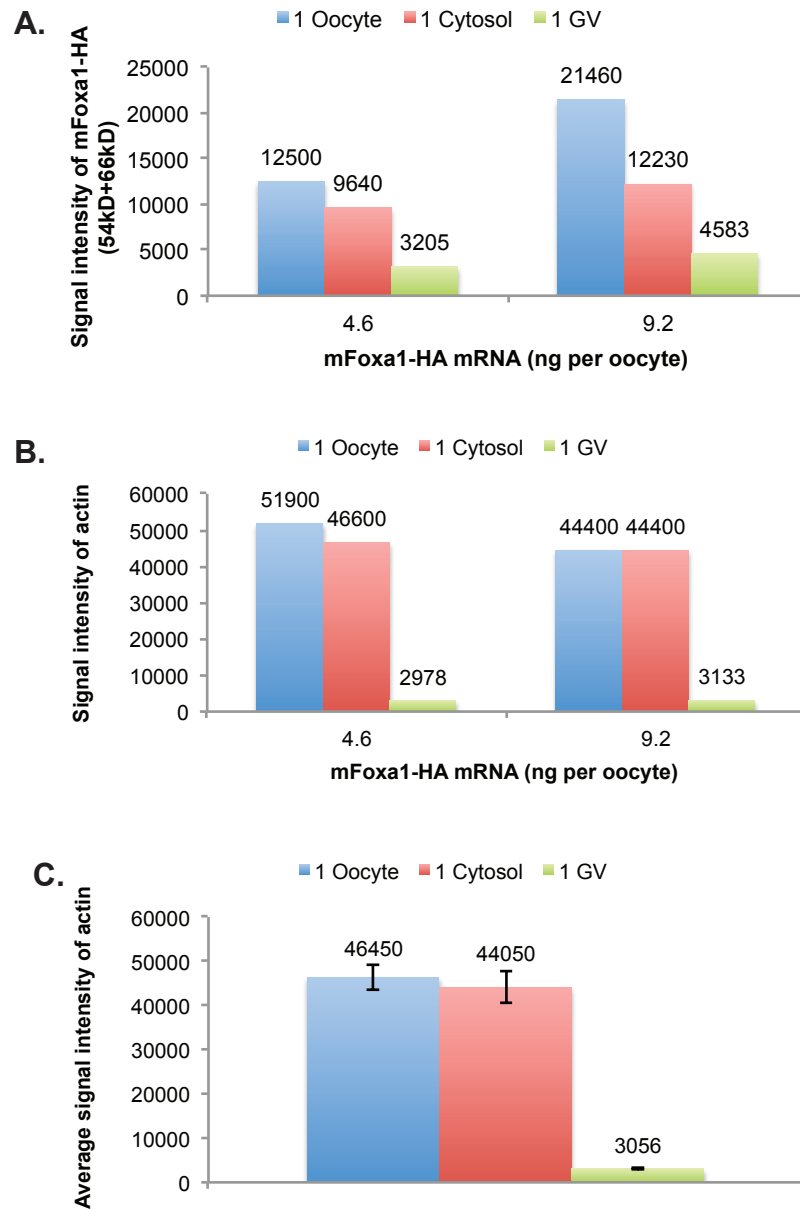


Figure 7.2.2 mFoxa1-HA proteins increase dose-dependently in GV and cytosol at Day 1 after mRNA injection.

- (A) mFoxa1-HA protein signals in different cellular compartments increase dose-dependently.
- (B) Different doses of mFoxa1-HA mRNA do not affect actin signals in different cellular compartments.
- (C) Average actin signals of 1 whole oocyte equal to average actin signals of 1 cytosol fraction plus 1 GV fraction.

Figure 7.2.3

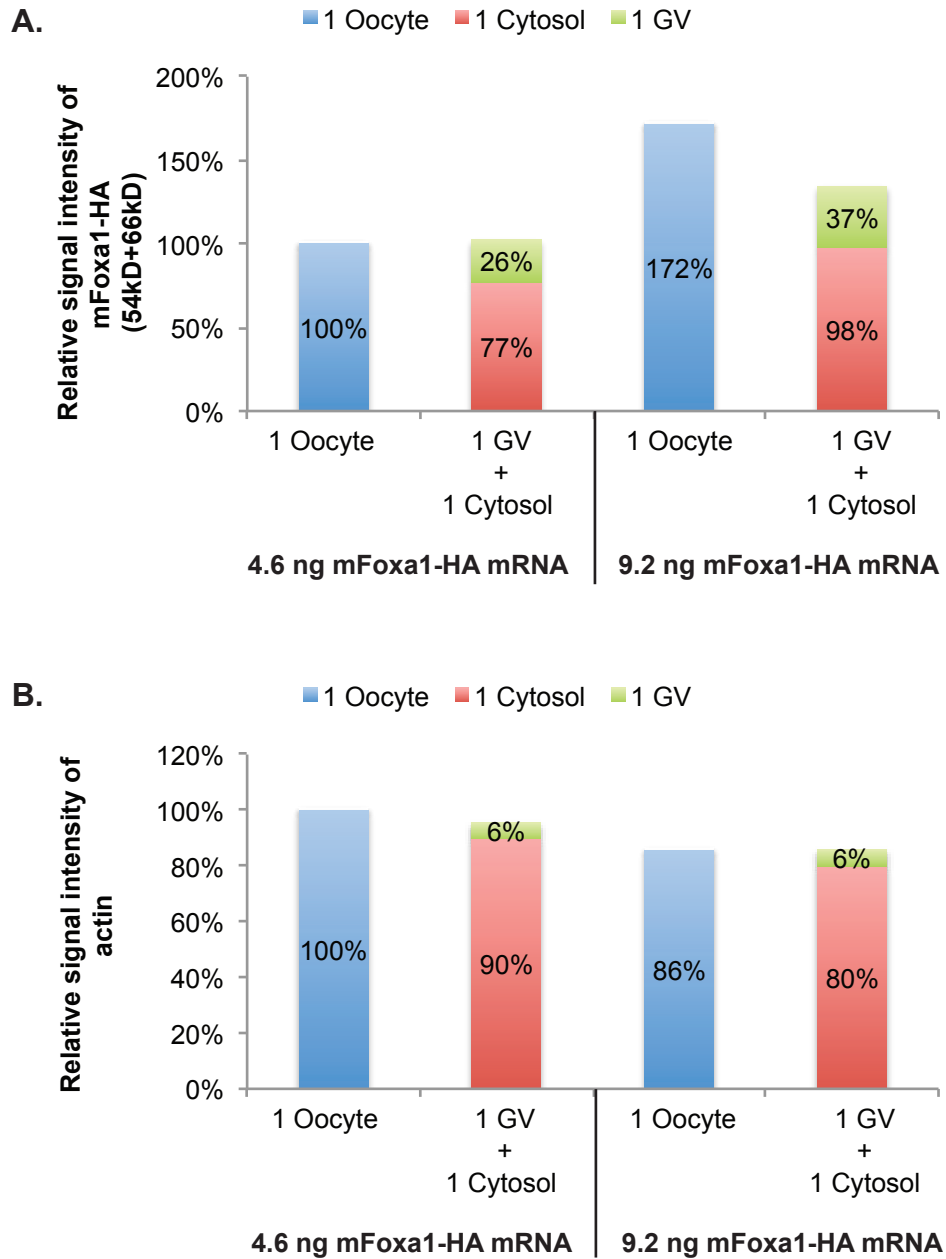


Figure 7.2.3 mFoxa1-HA proteins accumulate in GV dose-dependently at Day 1 after mRNA injection.

(A) About 1/4 of mFoxa1-HA proteins in 1 whole oocyte is distributed to 1 GV fraction regardless of the doses of mFoxa1-HA mRNA.

(B) About 1/15 of actin proteins in 1 whole oocyte is in 1 GV fraction.

Figure 7.2.4

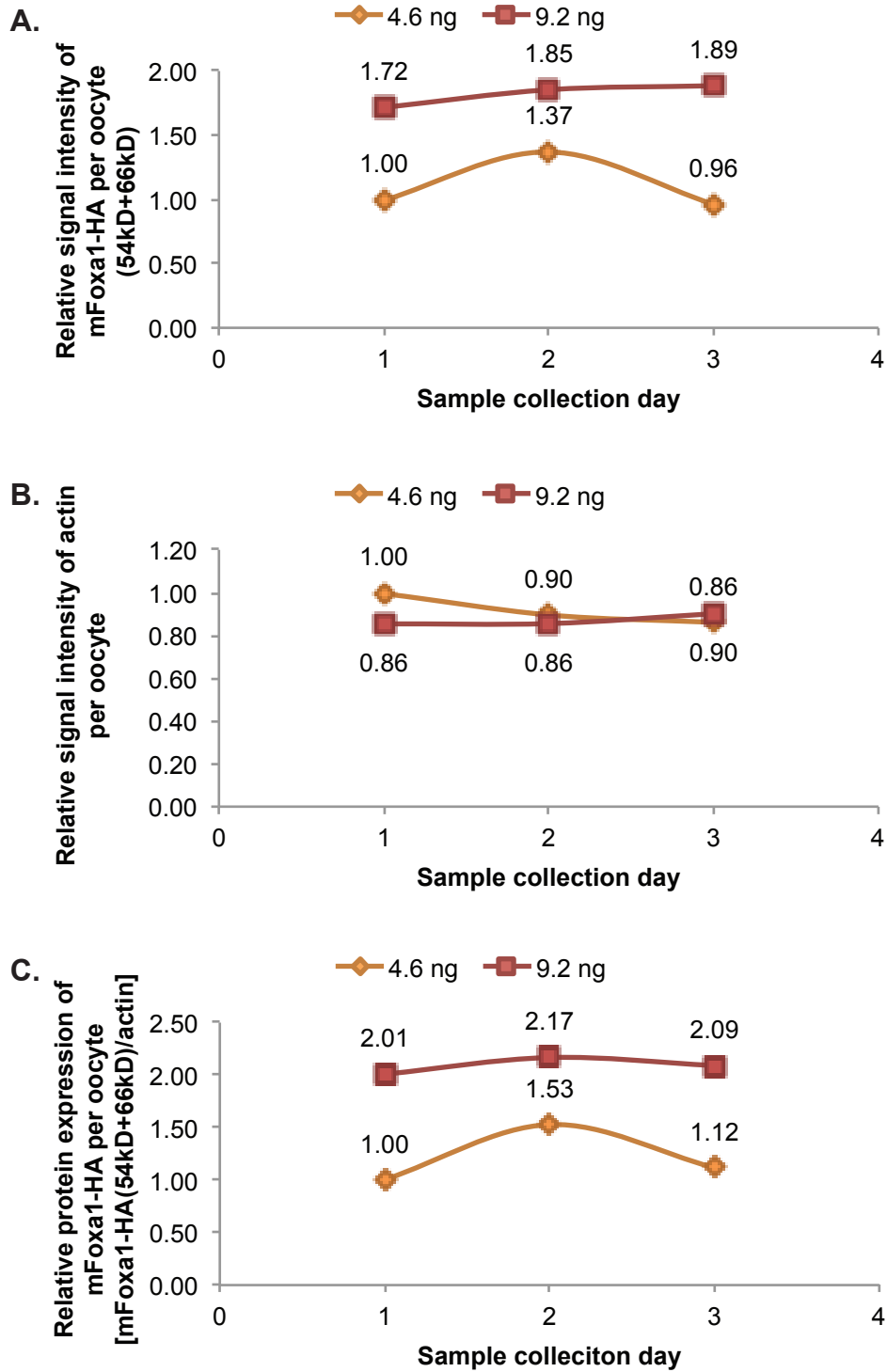


Figure 7.2.4 mFoxa1-HA proteins increase until Day 2 after mRNA injection and decrease slightly afterwards.

(A) Relative signal intensity of mFoxa1-HA per oocyte at Day 1, 2 and 3 after mRNA injection

(B) Relative signal intensity of actin per oocyte at Day 1, 2 and 3 after mRNA injection

(C) Relative signal intensity of mFoxa1-HA per oocyte is normalized by relative signal intensity of actin and is shown as relative protein expression.

7.2.2 SUMOylation is observed in mFoxa1-HA protein detection

For the double bands detected by anti-HA antibody (Figure 7.2.1), it might be similar to xklf2-HA that the anti-HA antibody sometimes detects small-sized bands of proteins when the protease inhibitors are insufficient to inhibit the protein degradation during sample handling or when the protein degradation is part of the post-translation modification for maintaining the half-life of protein of interest (Chapter 3). However, it is different between mFoxa1-HA and xklf2-HA since that the small-sized bands of xklf2-HA only exist in the cytosol fractions but two detected bands of mFoxa1-HA exist both in the GV and cytosol fractions. The absence of small-sized band of xklf2-HA in the GV fractions indicates the small-sized bands result from protein degradation in the cytosols, rather than in the GVs.

A possible explanation for the unexpected size of the presumed mFoxa1-HA bands in Figure 7.2.1 is the SUMOylation of mFoxa1-HA, which results in a bigger size band, and the SUMOylation is quite often seen in the Fox family^{139,152}. It has been shown that the SUMOylation of mFoxa1 is related to nuclear mobility, transcriptional activity and chromatin occupancy¹³⁹. If the double bands are caused by SUMOylation, the large-sized bands would represent SUMOylated mFoxa1-HA proteins (black arrowhead, ~66kD = 54kD for mFoxa1-HA + 12kD for SUMO protein) and the small-sized bands would be the native mFoxa1-HA (black arrow, 54kD) (Figure 7.2.1). It is interesting that the amount of 66kD SUMOylated mFoxa1-HA is more than the amount of 55kD native mFoxa1-HA in all samples regardless of the fractions, time points or doses (Figure 7.2.5).

For the whole oocyte samples (Figure 7.2.5.A), the relative expression of SUMOylated mFoxa1-HA (66kD) to native mFoxa1-HA (54kD) is different according to the doses of mFoxa1-HA mRNA and sample collection days. At Day 1, the SUMOylated mFoxa1-HA is 1.67 times more than the native mFoxa1-HA at Day 1 after mRNA injection. Since it has been proved that the distribution of the mFoxa1-HA proteins to the GVs is the same in both 4.6 and 9.2 ng mRNA groups (Figure 7.2.3.A), the different relative expression of SUMOylated mFoxa1-HA to native mFoxa1-HA is not related to the mFoxa1-HA distribution. Additionally, the difference of relative expression is also not related the dose of mFoxa1-HA mRNA because the relative expression of the 4.6 ng mRNA groups (1.67X) is more than the relative expression of the 9.2 ng mRNA groups (1.34X) at Day 1 (Figure 7.2.5.A). Instead, the relative expressions show that the SUMOylated mFoxa1-HA proteins increase in a time-dependent way and more mFoxa1-HA proteins are SUMOylated when incubation time increases (Figure 7.2.5.A).

On the other hand, when analyzing the GV and cytosol samples (Figure 7.2.5.B and 7.2.5.C), the relative expressions of SUMOylated mFoxa1-HA to native mFoxa1-HA only increase slightly in GV and in cytosol samples. For the GV samples, the relative expressions increase from 1.30X in 4.6 ng mRNA groups to 1.46X in 9.2 ng mRNA groups at Day 1 after mRNA injection. For the cytosol samples, the relative expression increase from 1.13X in 4.6 ng mRNA groups to 1.18X in 9.2 ng mRNA groups at Day 1 after mRNA injection. However, if comparing the difference of mFoxa1-HA SUMOylation between

GV samples and cytosol samples, the relative expressions of mFoxa1-HA SUMOylation is 1.38X in GV samples, which is more than the 1.16X in cytosol samples, and it might imply some functions involved SUMOylation and may relate to transcriptional activity of mFoxa1¹⁵².

Figure 7.2.5

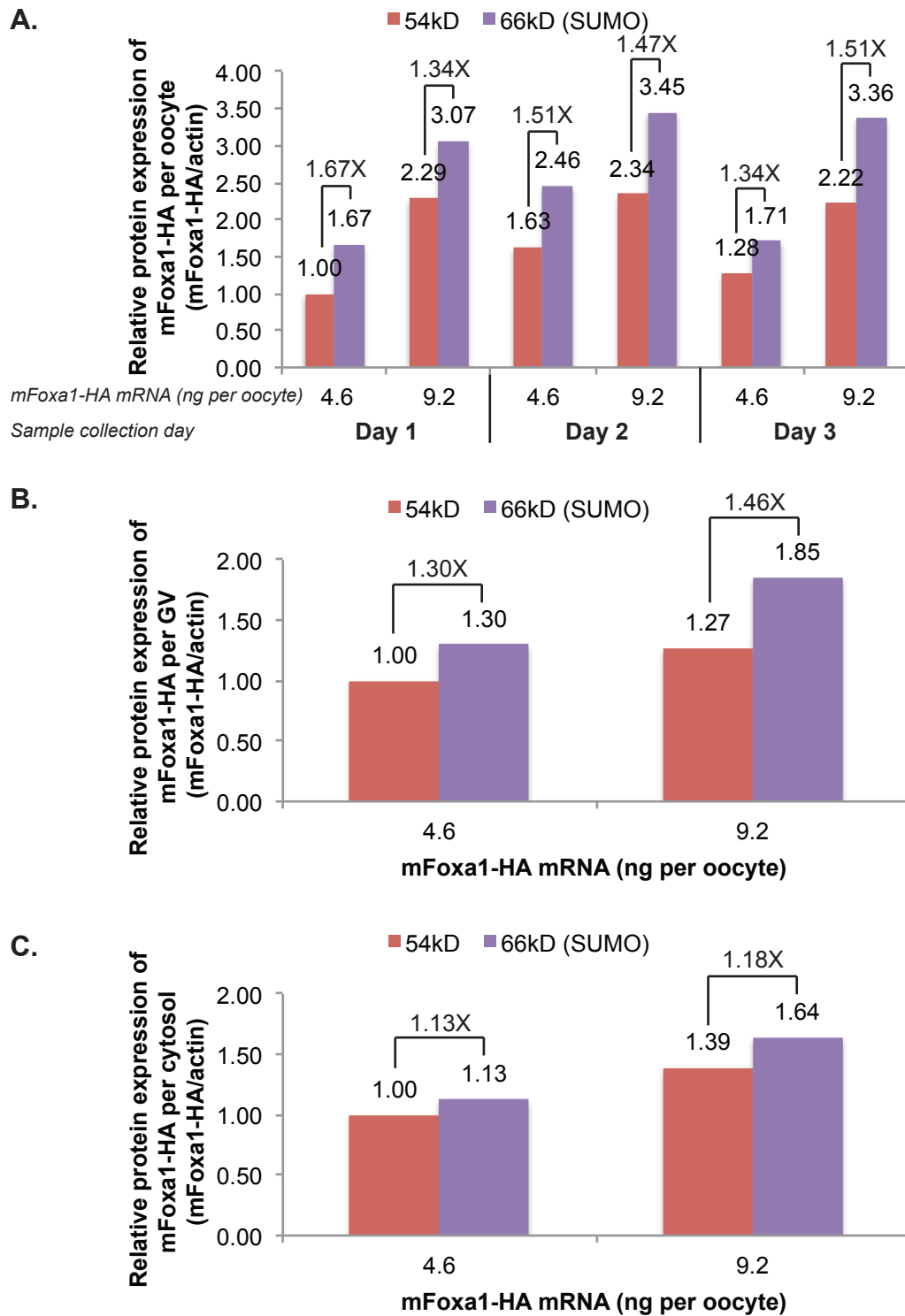


Figure 7.2.5 More mFxa1-HA are SUMOylated dependent on the increase of incubation time.

(A) The relative expressions of SUMOylated mFxa1-HA to native mFxa1-HA for oocyte samples at different time points and doses are shown.

(B and C) The relative expressions of SUMOylated mFxa1-HA to native mFxa1-HA for GV samples (B) and cytosol samples (C) at Day 1 after mRNA injection.

7.2.3 Summary

In all, this section proved that mFoxa1-HA proteins accumulate in the GVs dose-dependently after being translated from mRNA in the cytosol and the distribution of mFoxa1-HA proteins in the GVs does not change when introducing more mRNA into the cytosol. Additionally, the amount of mFoxa1-HA proteins increases in a dose- and time-dependent manner in the whole oocyte samples and decreases slightly at Day 2 after mRNA injection. Interestingly, slightly increased SUMOylation is observed on the newly-made mFoxa1-HA proteins and it might relate to the functions of mFoxa1-HA^{139,152}.

In conclusion, the fractionation of *Xenopus* oocytes provides a way to interpret the functions of proteins of interest further via a quantitative image system. Moreover, the translational machinery of *Xenopus* oocytes is able to translate a lineage-specification transcription factor, mFoxa1-HA, and process it constantly with SUMOylation, which can be further validated by anti-SUMO antibodies. This oocyte mRNA injection procedure therefore seems to provide a satisfactory assay to access the effect of mFoxa1-HA on transplanted somatic cell nuclei.

7.3 The overexpression of mFoxa1-HA up-regulates mSox2 only in MEFs during SCNR by *Xenopus* oocytes

In Chapter 3, it has been shown that each Yamanaka factor can selectively up-regulate tested pluripotency genes. In Chapter 4, it has also been shown that oocyte factors can up-regulate tested pluripotency genes up to ~600-fold while overexpression of xklf2-HA has smaller effect on tested pluripotency genes. Here, I used mFoxa1-HA, a lineage-specification factor and pioneer transcription factor, to examine if overexpression of mFoxa1-HA affects the expression of pluripotency genes during SCNR by *Xenopus* oocytes.

7.3.1 The overexpression of mFoxa1-HA up-regulates mSox2 among tested pluripotency genes

To evaluate the effect of mFoxa1-HA overexpression on tested pluripotency genes, Oocyte-NT (MEF cell line, sixiFM) is performed one day after mRNA injection and samples are collected 48 hours after Oocyte-NT (Figure 7.1.B). In Figure 7.3.1, I statistically analyzed the relative expression of tested pluripotency genes in MEFs, namely the genes, mSox2, mKlf4, mSall4, mOct4, mUtf1, mFbxo15, mNanog, mMyc, mEsrrb and mKlf2 (n=3, *t*-test).

Among the tested pluripotency genes, mSox2 is up-regulated by mFoxa1-HA overexpression with relative expression of 4.1, compared to oocyte factors (No mRNA injection groups, Figure 7.3.1). Other tested pluripotency genes are unaffected by mFoxa1-HA overexpression with relative expression less than 2 although mKlf4, mUtf1 and mFbxo15 are shown to be mildly up-regulated by mFoxa1-HA overexpression significantly (relative expression

less than 2, $p<0.05$, Figure 7.3.1). It is hard to judge if a transcription factor regulates the downstream genes with relative expressions close to one although the relative expressions are statically meaningful.

Figure 7.3.1

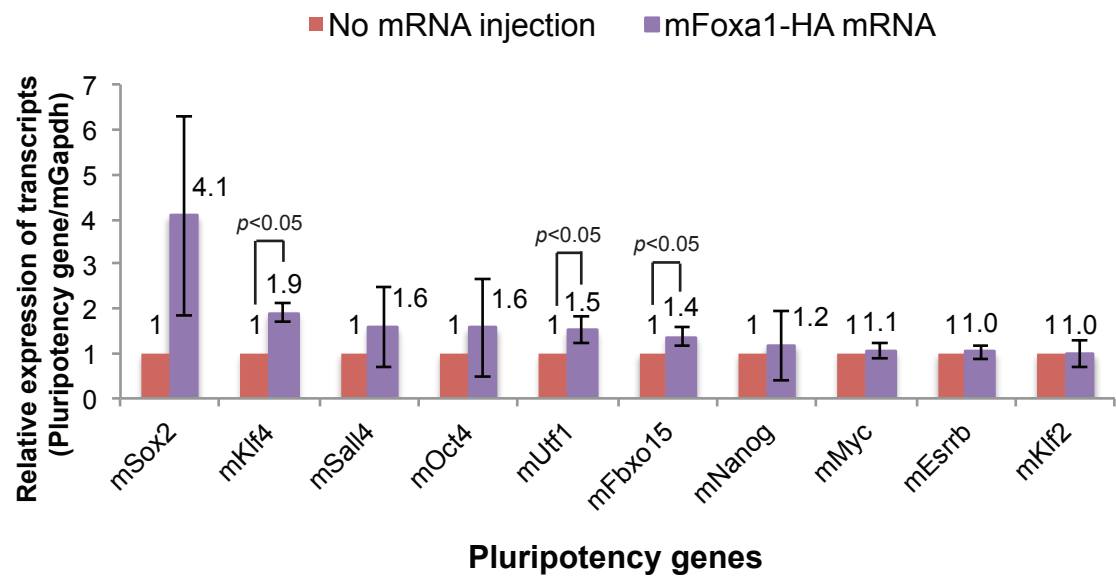


Figure 7.3.1 Among ten tested pluripotency genes, the overexpression of mFoxa1-HA up-regulates expression of only mSox2 in sixiFM MEF at Day 2 after Oocyte-NT.

7.3.2 The overexpression of mFoxa1-HA up-regulates tested pluripotency genes with variable relative expression in each biological replicate

In Figure 7.3.1, standard deviation of some genes is huge in the case of mSox2 (4.1-fold, $p < 0.08$), mOct4 (1.6-fold, $p < 0.45$) and mSall4 (1.6-fold, $p < 0.35$) and this may eliminate their biological meaning. Therefore, I compared the relative expressions of each gene from different frogs and examined how the individual difference affects the data readout (Figure 7.3.2). The frogs listed in Figure 7.3.2 are the same as those used in Figure 7.3.1 for statistical analysis.

For mSox2, the relative expressions of mSox2 are more than two in Frog 1, 2 and 3 without statistical significance (Relative expression=3, 2.6 and 6.6, Figure 7.3.2.A). For mSall4, the highest relative expressions are in the Frog 1 (2.6-fold), which are 2.5 times more than the lowest relative expressions in the Frog 3 (0.9-fold, Figure 7.3.2.C). For mOct4, the highest relative expressions are in the Frog 2 (2.8-fold), which are 3 times more than the lowest relative expressions in the Frog 1 (0.9-fold, Figure 7.3.2.D). Therefore, the statistical significance is difficult to judge owing to the individual differences among *Xenopus* oocytes.

Figure 7.3.2

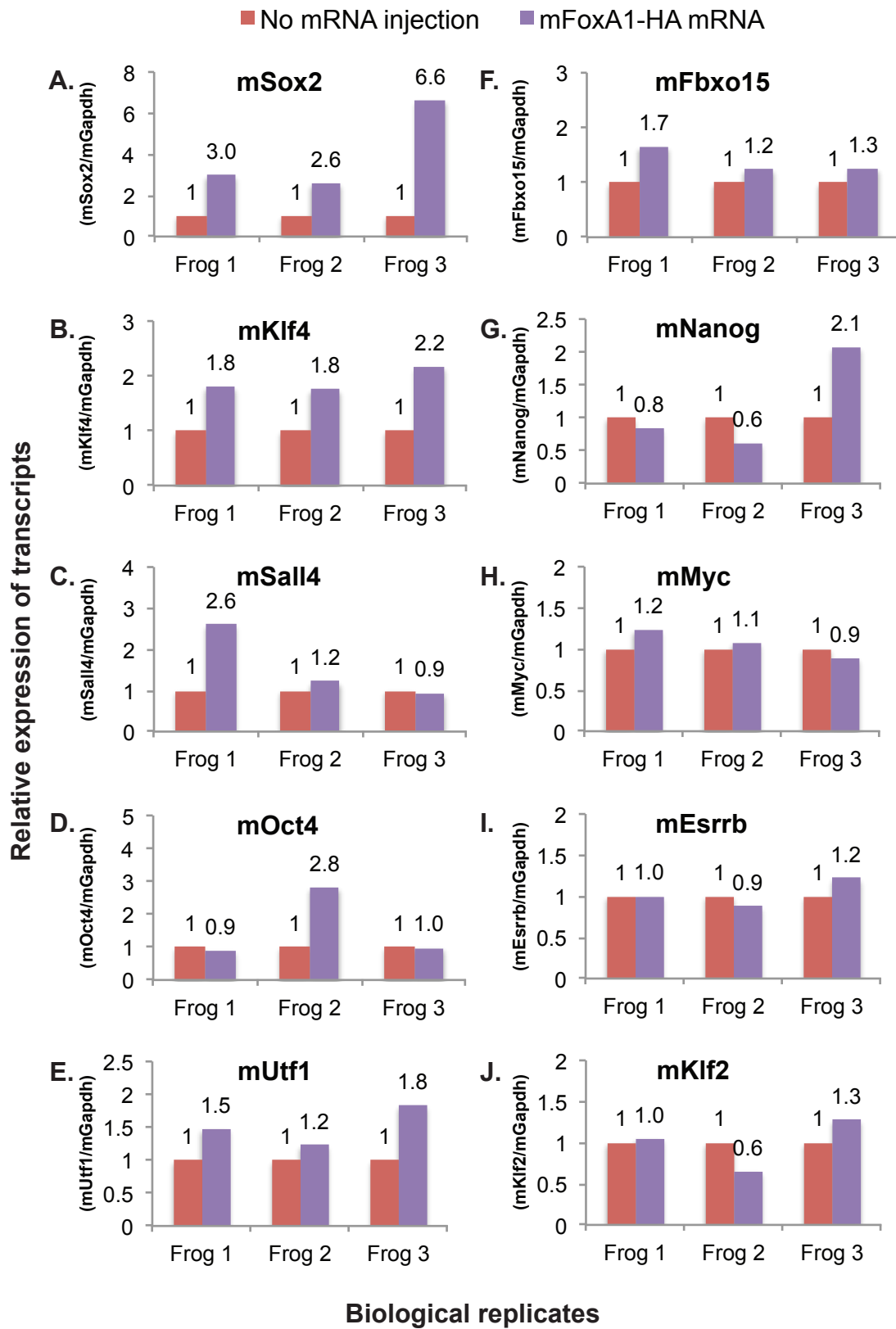


Figure 7.3.2 The overexpression of mFoxa1-HA up-regulates tested pluripotency genes with variable relative expression in each biological replicates.

7.3.3 Summary

To sum, the overexpression of mFoxa1-HA in the *Xenopus* oocytes mildly up-regulates only mSox2 among ten tested pluripotency genes. Combining both statistic analysis at Day 2 after Oocyte-NT and the change of transcripts at different time points after Oocyte-NT, mFoxa1-HA overexpression can up-regulate mSox2 during SCNR by oocytes even though mFoxa1, as a lineage-specification factor, may not exist or be important in early developmental stages before MBT.

7.4 The regulation of neurogenic genes by mFoxa1 is different between Oocyte-NT and neurogenesis of mDA neurons/progenitors

mFoxa1 is well characterized as a pioneer transcription factor, and can access low signal chromatin (the intermediate chromatin state between the active and repressed chromatin) and promote the recruitment of other transcription factors to regulate downstream genes¹⁴⁹. In addition, the early expression of mFoxa1 during normal neurogenesis as well as the effect of mFoxa1/2 on the neurogenic genes bound by mFoxa1 are recognized as *in vivo* evidence for mFoxa1 utilizing pioneer factor activities to affect the specification and differentiation of midbrain dopaminergic (mDA) neurons/progenitors *in vivo*^{150,151}. With these pioneer factor capabilities during neurogenesis, I wonder how mFoxa1 regulates the neurogenic genes in the context of somatic cell nuclear reprogramming by *Xenopus* oocytes and how effective it is to regulate neurogenic genes when compared with the oocyte factors in *Xenopus* oocytes?

In this section, Oocyte-NT (MEF, sixiFM) is performed one day after the injection of mFoxa1-HA mRNA and samples are collected on either Day 0 or Day 2 after Oocyte-NT (Figure 7.1.B). The neurogenic genes are selected from the paper described previously and are bound by mFoxa1 and mFoxa2 during the differentiation of mDA neurons/progenitors on mouse embryonic day (E) 12.5¹⁵⁰. In the referenced paper, mFoxa1 and mFoxa2 function redundantly during the specification and differentiation of mDA neurons/progenitors and a mFoxa1/2 double cKO mutant is made for

observing how the mFoxa1/2 binds and regulates the neurogenic genes (cko, conditional knockout, Cre-Lox system)¹⁵⁰.

Table 7.4 shows how the mFoxa1/2-bound neurogenic genes are regulated in the mFoxa1/2 cko embryos¹⁵⁰. In the first category, Arx, Dmrtb1, Ferd3l, Kcnip3, Lmcd1, Lmx1b, Rora and Smarca1 are down-regulated in the mFoxa1/2 cko embryos and it suggests they may be up-regulated by mFoxa1/2 in normal mDA neurons/progenitors on E12.5. The second category includes Gli1, Meis2, Otx1 and Otx2, which are up-regulated in the Foxa1/2 cko embryos and therefore may possibly be down-regulated in the normal mDA neurons/progenitors on E12.5. The third category shows Elk3 and Foxb1 are not affected in the mFoxa1/2 cKO mDA neurons/progenitors on E12.5 although they are bound by mFoxa1/2. From the in vivo data of mFoxa1/2 bound genes during mDA neurons/progenitors differentiation (Table 7.4), I now ask how these genes are regulated in *Xenopus* oocytes where the oocyte factors might differ hugely from the endogenous factors of mDA neurons/progenitors.

Table 7.4

<u>Down-regulated mFoxa1 ChIP target genes in the Foxa1/2 cKO embryos</u>				
Arx	Dmrtb1	Ferd3l	Kcnip3	Lmcd1
Lmx1b	Rora	Smarca1		
<u>Up-regulated mFoxa1 ChIP target genes in the mFoxa1/2 cKO embryos</u>				
Gli1	Meis2	Otx1	Otx2	
<u>Un-affected mFoxa1 ChIP target genes in mFoxa1/2 cKO embryos</u>				
Elk3	Foxb1			

Table 7.4 Regulation of mFoxa1 ChIP target genes in the mFoxa1/2 cKO embryos.

Genes are identified by ChIP-seq and RNA-seq.

7.4.1 The overexpression of mFoxa1-HA up-regulates mDmrtb1, mFoxb1 and mOtx2 in sixiFM MEF up to 31-fold at Day 2 after Oocyte-NT

In Figure 7.4.1, the mFoxa1-HA mRNA is injected into *Xenopus* oocytes one day before Oocyte-NT and the samples are collected at Day 2 after Oocyte-NT (n=3, Figure 7.1.B, page 247). Fourteen neurogenic genes are evaluated here and analyzed by QPCR. Compared to the no mRNA injection groups, three neurogenic genes are up-regulated by the overexpression of mFoxa1-HA (mDmrtb1, 2.3-fold; mFoxb1, 4.3-fold; mOtx2, 31-fold, Figure 7.4.1.A). While mDmrtb1 is down-regulated, mOtx2 is up-regulated and mFoxb1 is unaffected in the mFoxa1/2 cKO embryos (Table 7.4), the overexpression of mFoxa1-HA has a different effect on these genes during SCNR by oocytes.

On the other hand, eleven neurogenic genes are unaffected by the overexpression of mFoxa1-HA with relative expressions less than two and without statistical significance ($p>0.05$, Figure 7.4.1.B). Within these unaffected genes, five genes have relative expression less than one (Figure 7.4.1.B: mLmcd1, mFerd3l, mKcnip3 and mElk3) and the relative expressions of Lmcd1 and Ferd3l are not determined because the relative amount of transcripts is under the detection limit of QPCR machine (Figure 7.4.1.B). It is an indication for that these genes are silent in the sixiFM MEF. The other gene sets among these unaffected genes with relative expression more than one but less than two are mLmx1b, mOtx1, mSmarca1, mRora, mMeis2, mArx and mGli1 (Figure 7.4.1.B).

Figure 7.4.1

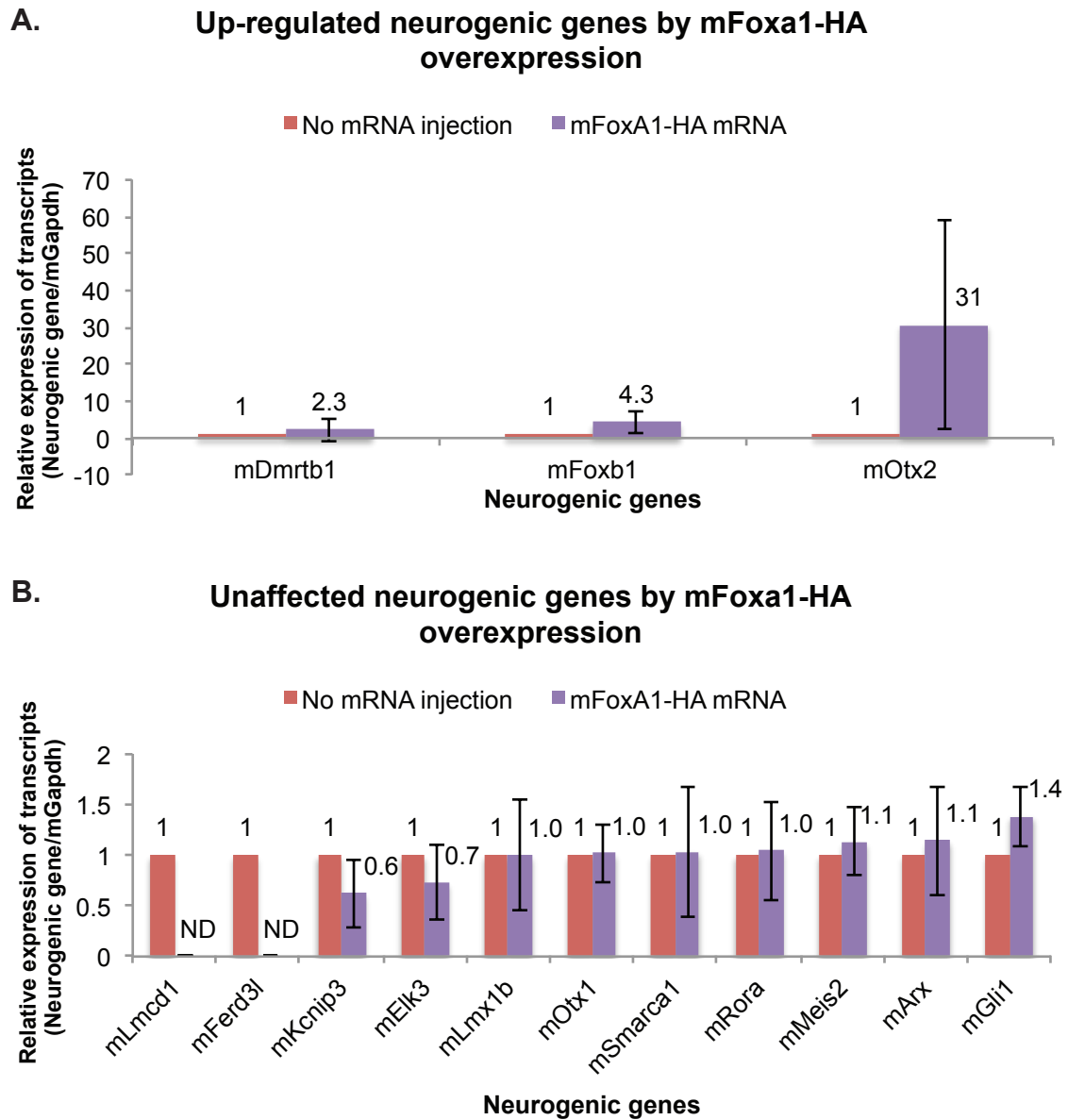


Figure 7.4.1 Effect of mFoxa1-HA overexpression on neurogenic genes at Day 2 after Oocyte-NT is shown (n=3, *t*-test).

Neurogenic genes, which are ChIP targets of mFoxa1 in mDA neurons/progenitors, are up-regulated (A) and unaffected (B) by mFoxa1-HA overexpression. None of these genes are shown to be significantly regulated by mFoxa1-HA overexpression. ND, not determined.

7.4.2 Oocyte factors up-regulate mOtx2, mFerd3l, mKcnip3, mOtx1, mGli1 in sixiFM MEF up to 313-fold from Day 0 to Day 2 after Oocyte-NT

To compare the effect of oocyte factors and mFoxa1-HA on neurogenic genes, the samples are collected at Day 0 and Day 2 after Oocyte-NT (Figure 7.1.B, page 247). Samples for the time-point observation are the same as the samples of Frog 1 in Figures 7.4.4 and 7.4.5.

For the up-regulated neurogenic genes, mDmrtb1, mFoxb1 and mOtx2 have been shown to be up-regulated by mFoxa1-HA overexpression at Day 2 after Oocyte-NT (Figure 7.4.1.A). Although these three genes are up-regulated by mFoxa1-HA overexpression, the oocyte factors affect them differently (Figure 7.4.2). This shows that mOtx2 is up-regulated by oocyte factors from Day 0 to Day 2 after Oocyte-NT (1-fold to 17-fold, Figure 7.4.2.C) whereas mDmrtb1 and mFoxb1 are unaffected by oocyte factors (mDmrtb1, 1-fold to 1.2 fold; mFoxb1, 1-fold to 0.9-fold, Figure 7.4.2.A and B). The unmatched gene regulation may be caused by the different contents between *Xenopus* oocytes at stage V/VI and mDA neuron/progenitors on E12.5 and different chromatin states between sixiFM MEF and mDA neurons/progenitors.

For other neurogenic genes, mLmcd1, mFerd3l, mKcnip3, mElk3, mLmx1b, mOtx1, mSmarca1, mRora, mMeis2, mArx and mGli1 have been shown to be unaffected by mFoxa1-HA overexpression at Day 2 after Oocyte-NT (Figure 7.4.1.B). Among these genes, mFerd3l, mKcnip3, mOtx1, mGli1 are up-regulated by oocyte factors from Day 0 to Day 2 after Oocyte-NT with relative expression from 8.2 to 313 (Figure 7.4.3.B, C, F and K). Some data of

mFerd3l and mRora are undetermined due to the nil or low expression level of these two genes in sixiFM MEF (Figure 7.4.3.B and H).

Figure 7.4.2

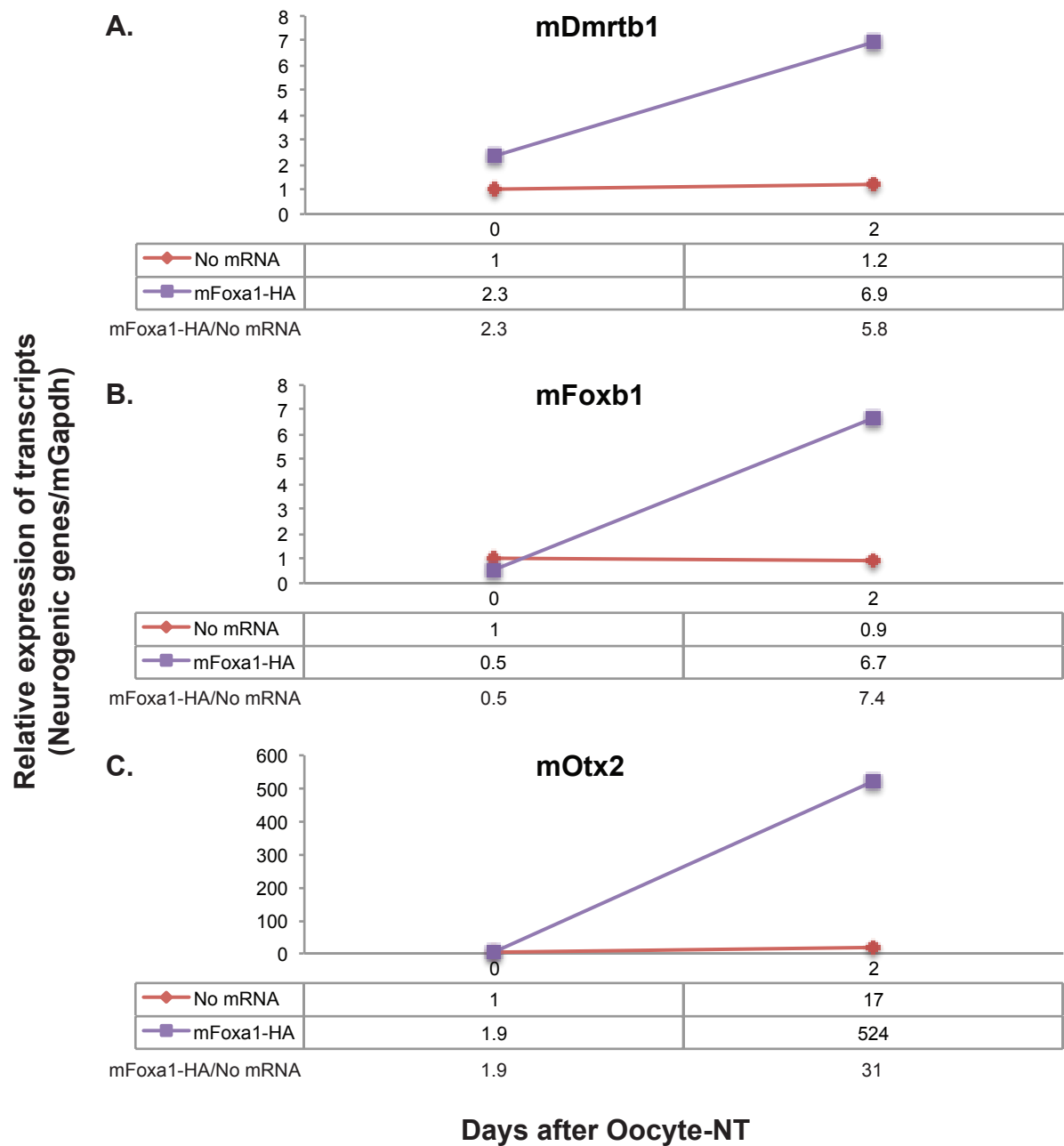


Figure 7.4.2 In Frog 1, mOtx2 (C) is up-regulated by oocyte factors and mFoxa1-HA overexpression while mDmrtb1 (A) and mFoxb1 (B) are up-regulated by mFoxa1-HA overexpression but not affected by maternal factors due to the SCNR resistance.

Figure 7.4.3

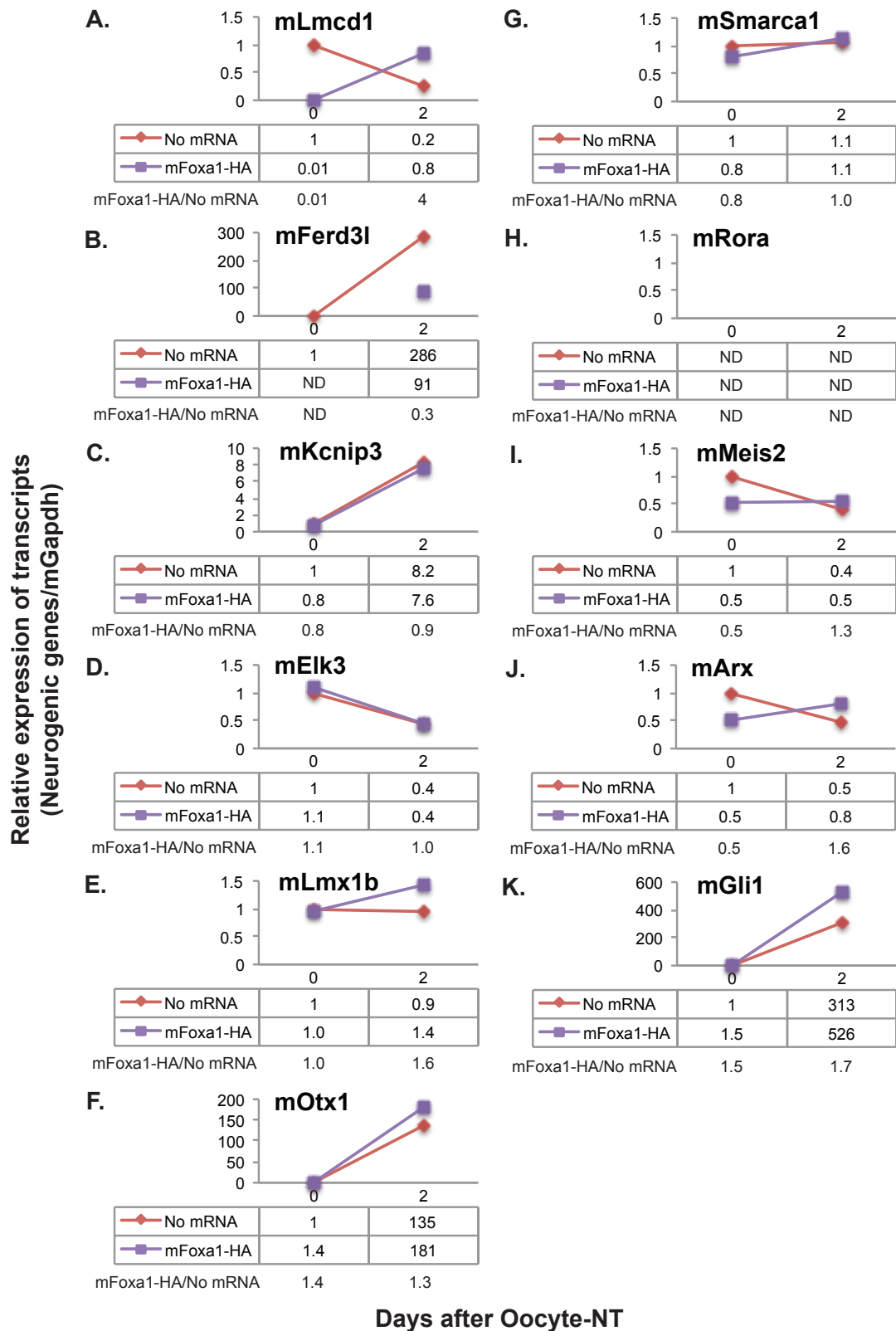


Figure 7.4.3 While mFerd3l, mKcnip3, mOtx1 and mGli1 are strongly up-regulated by oocyte factors from Day 0 to Day 2 after Oocyte-NT, all these neurogenic genes are unaffected by mFoxa1-HA overexpression at Day 2 after Oocyte-NT.

7.4.3 mFoxa1-HA overexpression up-regulates mDmrtb1, mFoxb1 and mOtx2 in sixiFM MEF with variable relative expression due to SCNR resistance

In Figure 7.4.4, the relative expression of three up-regulated genes, mDmrtb1, mFoxb1 and mOtx2, are shown for three different frogs (Figure 7.4.4). For mDmrtb1, expression is up-regulated by 5.9-fold in Frog 1 but it seems to be down-regulated in Frog 2 and 3 with relative expression less than 1 (Figure 7.4.4.A). The opposite regulation of mDmrtb1 might be caused by the nil or low expression of it in sixiFM MEF and mDmrtb1 in some samples is successfully activated by mFoxa1-HA overexpression but some are not.

For mFoxb1 and mOtx2, they are clearly up-regulated by mFoxa1-HA overexpression in all three frogs although the relative expressions are variable (Figure 7.4.4.B and C). As mentioned previously for the effect of xklf2-HA overexpression, xklf2-HA overexpression up-regulates target genes with variable and high relative expression when target genes are resistant to oocyte factors. Similarly, mFoxa1-HA overexpression is shown to have the same effect on up-regulating target genes, which are resistant to SCNR by oocyte factors.

As for the unaffected genes, mLmcd1 are up-regulated by mFoxa1-HA overexpression by 3.4-fold in Frog 1 and mFerd3l are up-regulated by 24-fold in Frog 2 (Figure 7.4.5.A and B). Although these two genes are not judged to be up-regulated by mFoxa1-HA overexpression, they are possible mFoxa1-HA target genes and the random activation by mFoxa1-HA overexpression is caused by the resistance of mLmcd1 and mFerd3l to oocyte factors.

Figure 7.4.4

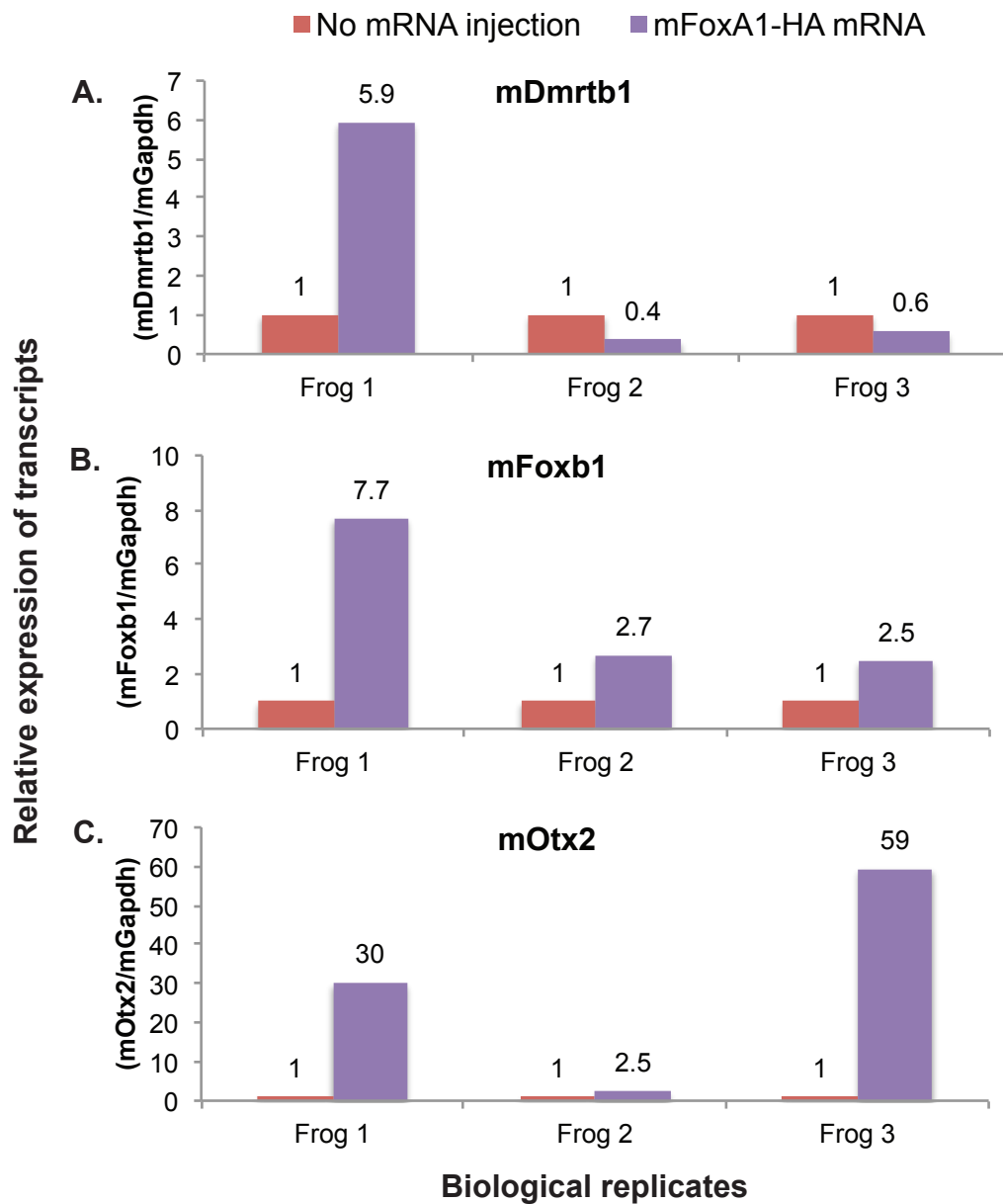


Figure 7.4.4 mDmrtb1, mFoxb1 and mOtx2 are up-regulated by mFox1-HA overexpression with variable relative expression in different frogs due to SCNR resistance of these genes in sixiFM MEF at Day 2 after Oocyte-NT.

Figure 7.4.5

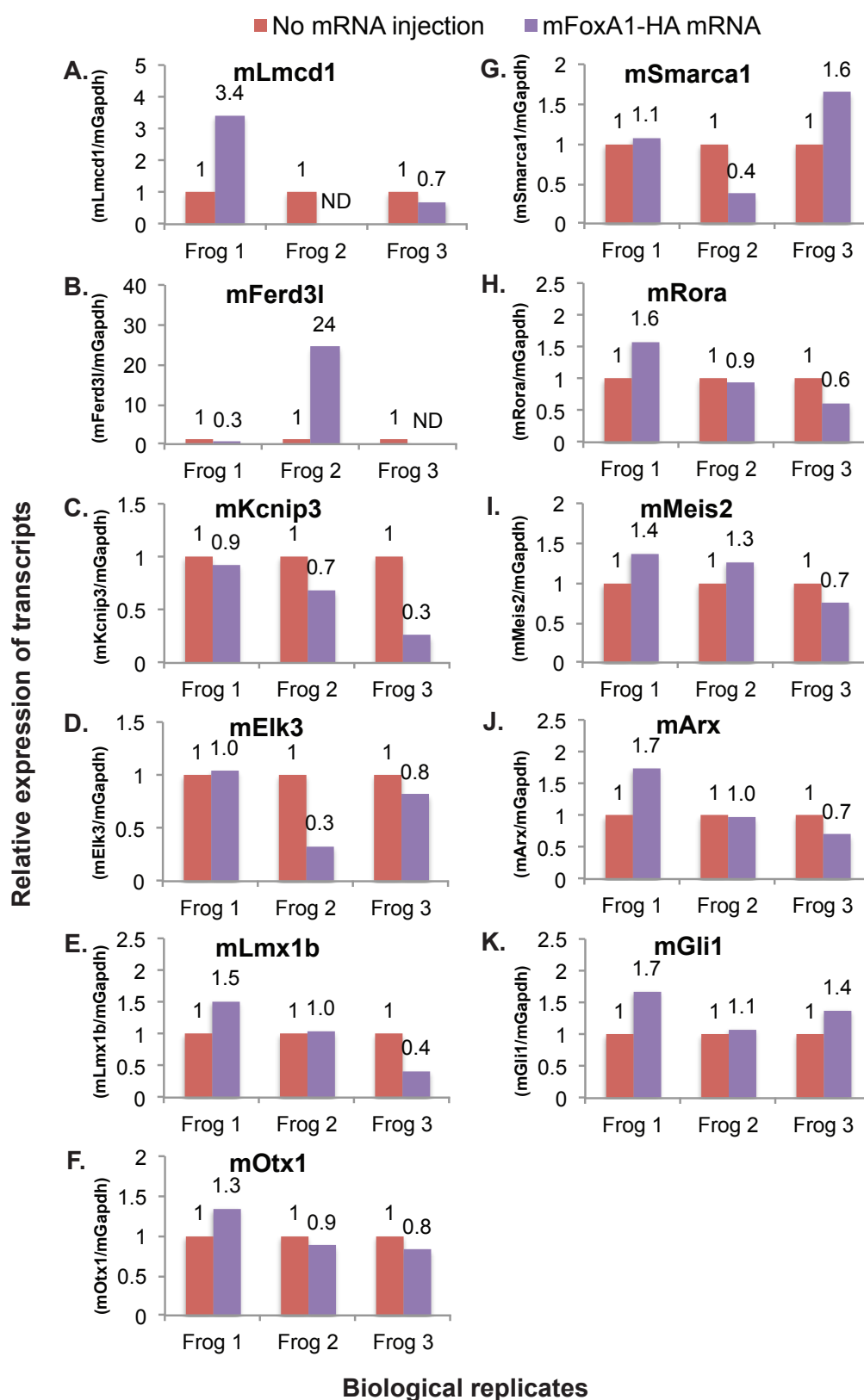


Figure 7.4.5 Among mFoxa1-HA unaffected genes, mLmcd1 and mFerd3l are possible mFoxa1-HA target genes while they are up-regulated by mFoxa1-HA overexpression in some Frogs but not in other Frogs due to SCNR resistance.

7.4.4 Summary

The overexpression of mFoxa1-HA up-regulates three neurogenic genes, mDmrtb1, mFoxb1 and mOtx2 and does not affect other eleven neurogenic genes at Day 2 after Oocyte-NT. Although these neurogenic genes are bound by mFoxa1/2 in the mDA neurons/progenitors during neurogenesis, they are regulated by the expression of mFoxa1 in the mDA neurons/progenitors differently from being regulated during Oocyte-NT. All in all, mFoxa1-HA overexpression up-regulates some target genes with high and variable relative expression due to the SCNR resistance of these target genes.

7.5 Conclusions

In the early developmental processes of *Xenopus laevis*, relative expression of *xfoxa1* is activated after MBT and maintained at the same level until stage 33^{140,141}. Additionally, mFoxa1 has been shown to be crucial for the induction of neurogenic genes in mDA neurons/progenitors in neurogenesis^{150,151}. It seems *xfoxa1* or mFoxa1 is needed for later developmental processes, such as lineage specification, but not for the early developmental stages before MBT. To examine the role of mFoxa1 during SCNR by oocytes, I overexpressed mFoxa1-HA and examined how it affected pluripotency genes and neurogenic genes after Oocyte-NT.

First, I have shown that mFoxa1-HA protein can be produced from mFoxa1-HA mRNA injected into cytosols of *Xenopus* oocytes time- and dose-dependently and mFoxa1-HA proteins accumulate in the GVs after being produced in the cytosols from mRNA dose-dependently. Furthermore, SUMOylation is observed on newly-made mFoxa1-HA proteins.

In terms of gene regulation, mFoxa1-HA overexpression has been shown to up-regulate only mSox2 in sixiFM MEF among ten selected pluripotency genes at Day 2 after Oocyte-NT. Therefore, mFoxa1-HA overexpression seems not to be important to induce pluripotency genes. For neurogenic gene regulation, mFoxa1-HA overexpression has been shown to up-regulate mDmrtb1, mFoxb1 and mOtx2 in sixiFM MEF among 14 neurogenic genes at Day 2 after Oocyte-NT. Interestingly, gene regulation of the same sets of neurogenic genes by mFoxa1-HA overexpression is different between in mDA

neurons/progenitors during neurogenesis and in *Xenopus* oocytes during Oocyte-NT.

In conclusion, the regulation of genes by overexpression of transcription factors is dependent on the factors in *Xenopus* oocytes and chromatin structure of donor nuclei during SCNR by oocytes. The cooperation between oocyte factors and transcription factors decides the regulation of transcription factor target genes. In addition to that, the overexpression of transcription factors can up-regulate target genes with small relative expression when the up-regulation is an enhancement of gene expression or with huge relative expression when the up-regulation is an activation of genes, more specifically, when these genes are resistant to activation by oocyte factors.

Chapter 8 Discussions

8.1 Introduction

The ultimate form of nuclear reprogramming is to derive a complete normal adult animal, and hence all cell-types, from the nucleus of a specialized cell. However, nuclear transfer to eggs and induced pluripotency are successful at a low frequency and this obstructs attempts to understand the mechanisms involved. In our lab we have made use of non-dividing oocytes (egg progenitors) to identify some steps in the nuclear reprogramming process. However, somatic nuclei injected into oocytes (amphibian) do not yield adult animals or even functional adult cells. Therefore, to what extent can somatic cell nuclear transfer to oocytes be used to analyse nuclear reprogramming?

Complete nuclear reprogramming requires two steps. The first is to rejuvenate gene expression in adult cells back to that of an embryo. This requires the activation of embryo-expressing genes such as Oct4 and Sox2, which are usually dormant in adult cells. The second requires the suppression of cell-type-specific genes that are active in adult cells. Nuclear transfer to oocytes achieves the first of these steps. It also achieves several of the clearly seen changes in chromatin configuration, including nuclear swelling and chromatin dispersion¹⁵³. A special advantage of nuclear transfer to amphibian oocytes is that the chromatin changes observed applies to almost all transplanted nuclei from about 300-500 cells injected into an oocyte GV¹⁵⁴. This cannot be said of

nuclear transfer to eggs or of induced pluripotency. We can therefore use nuclear transfer to oocytes to identify the first steps of reprogramming.

All of the donor nuclei used in the work described here are nuclei from mammalian cells. We have done this because the genome of mice is fully sequenced making it possible to identify genes whose expression is changed in these experiments. Using mouse cells to test response avoids the possible problem of distinguishing mouse transcripts from homologous *Xenopus* transcripts contained in the recipient *Xenopus* oocyte. Although cross-species nuclear transfer to eggs do not develop normally¹⁵⁵, much work with induced pluripotency has made use of cross-species combinations. Therefore an understanding of reprogramming mechanisms observed by nuclear transfer to oocytes (Oocyte-NT) can benefit from the cross-species interactions.

In this chapter I review the effects of endogenous factors and *xklf2* overexpression, and this is specifically on those genes that are believed to be important in pluripotency. Then I discuss the extent to which oocyte factors and overexpression of transcription factors activate reprogramming-resistant genes. Finally, I discuss possible mechanisms involved in nuclear reprogramming via nuclear transfer, induced pluripotency and cell fusion.

8.2 Gene reprogramming by maternal factors of *Xenopus* oocytes and by TF overexpression

In this thesis, I have used two different procedures, namely QPCR and RNA-seq plus BrUTP pull-down, to evaluate the effects of maternal factors and *xklf2*-HA overexpression on gene regulation and hence to measure gene resistance against SCNR.

The first method evaluates changes in the number of transcripts of individual genes over two days from the time of nuclear transfer (Day 0) to two days later (Day 2). For this, I have used QPCR (see Materials and methods, Chapter 2, Page 46). By this procedure I quantitate any change in the number of transcripts of a particular gene. The values obtained therefore reflect the balance between synthesis and turnover of transcripts for a gene. The values represent the steady state concentration in the oocyte nucleus for that transcript; they include any transcripts for that gene that have been carried over in the injected nuclei from the donor cell population. It does not therefore measure directly the transcriptional activity of a gene. It is this concentration of a gene product which is important in influencing other downstream genes whose activity is dependent on the gene in question.

Using the same assay (QPCR), I have also assessed the effect of overexpressing the Yamanaka factor, *xklf2*-HA. This factor was chosen because it was found to affect more pluripotency genes than the others, and because its expression level is higher in early embryos than that at later stages after MBT, suggesting its role as a maternal factor in oocytes.

The fold changes in gene expression obtained by QPCR are shown in Figure 8.2.1, 8.2.2, 8.2.3 and 8.2.4. For each figure, I have chosen to present the results for mSox2 and mOct4 separately from those for other genes because the values of fold change for these two genes are much higher than for the others. These figures obtained are also summarized in Tables 8.2.1, 8.2.2, 8.2.3, and 8.2.4. An overall summary is shown in Table 8.2.5.

I should first point out that the reference point for the values obtained are those for mGapdh, a well-expressed housekeeping gene, whose steady state concentration, measured by QPCR, does not change over the two-day period. Compared to this, the enhanced amount of mSox2 and mOct4 transcripts is enormous, ranging from over 10-fold to nearly 1500-fold for the two kinds of fibroblast cell lines tested. Changes for other genes are mild (4- to 23-fold) or not significant (0.5 to 1.9-fold) – see Table 8.2.1 and Summary Table 8.2.5.

If we look now at assays for the values of fold change for nuclei of mMyoblasts (C2C12), the results are broadly in line with what I have found for MEF cells. This is to say that the major effects are seen for mSox2 and mOct4 and a mild effect on mUtf1 and mJun; other genes show no effect (Figure 8.2.3). All these effects described reflect the activity of the oocyte's endogenous factors.

We can now look at the effect of overexpressing the Yamanaka factor *xklf2*, using now the *Xenopus* homologue. The results are shown in the figures just referred to as well as in Table 8.2.1. Table 8.2.5 lists the genes in decreasing

order of oocyte effect from strong (mSox2 and mOct4) to, mild, weak or nil (10 other genes). The striking result is that, in no case, does xkfl2 significantly affect what I have found for the strong effect (mSox2 and mOct4) by endogenous oocyte factors. It can be seen that in the column headed “xkfl2-HA overexpression” in Table 8.2.5 there is never a significant effect of xkfl2, whose effect is remarkably similar to the strong effect (mSox2 and mOct4) in the column “Oocyte factors”. This suggests that the endogenous oocyte factor, xkfl2 or its equivalent, is unlikely to be limiting in causing activation of the tested pluripotency genes by nuclear transfer. If it had been limiting, its overexpression should have had other effects.

It is particularly interesting to now compare the effect of overexpressing the pioneer factor, mFoxa1, in the present system. Figure 8.2.4 (small purple bars) shows that mFoxa1 has very little effect on mSox2 and mOct4 and no obvious effect on other pluripotency genes. Recall the results of mFoxa1 effect on neurogenic genes in Chapter 7 (Figure 7.4.1, page 269). Overexpressed mFoxa1 also cannot up-regulate most of these neurogenic genes in MEFs even these genes were proven to be its binding targets during development. It suggests that genes activated by overexpressed pioneer factors in Oocyte-NT not only require corresponding binding sites but cofactors while oocytes may not contain neurogenic factors that can cooperate with mFoxa1 for effective gene activation.

Overall, at least for the genes tested, and for the nuclei of the cell types tested, oocyte (endogenous) factors far outweigh the importance of the Yamanaka

factors and pioneer factors identified by an entirely different experimental system.

In this work, I have included a completely different assay system, RNA-seq plus BrUTP pull-down. As explained in the Materials and methods (Chapter 2, Page 47 and 48), this uses the incorporation of 5-bromo-UTP into RNA that is synthesized from the time of nuclear transfer for the next two days. As validated by Jullien et al¹³⁶, this quantitates the new synthesis of transcripts by each gene genome wide. This is a more accurate assessment of gene activation by oocytes compared to QPCR. In the first part of this section, I have looked at the steady state concentration of a transcript, and this includes any “carry over” RNA from the donor cells. In contrast, the BrUTP assay takes no account of carry over RNA. It represents the transcriptional activity of a gene from its time of introduction into an oocyte to the time of assay, two days later.

It is therefore a true measure of gene transcriptional activation. This difference accounts for the fact that the Ct values for mNanog and mLefty1 are close to Ct value of mGapdh and much lower than Ct of mSox2 and this indicates these two genes are considered highly expressed in donor cell types, MEFs and mMyos (QPCR, Table 8.2.1 and 8.2.3). After 2 days of nuclear transfer, the expression of these two genes is still as high as the level in donor cells before nuclear transfer (QPCR, Table 8.2.1 and 8.2.3). However, this high expression level of mNanog and mLefty1 (QPCR, Summary Table 8.2.6) is in contrast to the nil/low expression level reprogrammed by oocyte factors at

oocyte steady state in Figure 8.2.6.B, 8.2.6.C and Summary Table 8.2.7. Therefore, the amount of mNanog and mLefty1 transcripts, measured at Day 2 after Oocyte-NT by QPCR, are all from carry-over in donor cells and no new transcript synthesis induced by oocyte factors after nuclear transfer (RNA-seq plus BrUTP pull-down).

Considering this difference, the similar oocyte effects between data obtained from these two assays can therefore be seen in mESCs, MEFs and mMyos (Table 8.2.6 and 8.2.7). While mSox2 and mOct4 are strongly up-regulated by oocyte factors from Day 0 to Day 2 after Oocyte-NT in MEFs and mMyos (QPCR, Figure 8.2.3), the expression level of them at oocyte steady state is different and this difference shows the resistance of genes corresponding to certain cell types (RNA-seq plus BrUTP pull-down, Figure 8.2.5.A). In the case of mSox2, mSox2 in MEFs is reprogrammed to be highly expressed (FPKM>4) by oocyte factors at an oocyte-steady state but mSox2 in mMyos is reprogrammed to be lowly/mildly expressed (FPKM≤4) at an oocyte-steady state (Figure 8.2.5.A and Table 8.2.7). This is consistent to the results obtained by QPCR, which shows that mSox2 in MEFs is reprogrammed to highly expressed at Day 2 after Oocyte-NT and mSox2 in mMyos is reprogrammed to be lowly expressed, judged by the Ct value between mSox2 and mGapdh (QPCR, Table 8.2.3 and 8.2.6). For mOct4 in MEFs and mMyos, the results are also consistent by using both assays and both results show that mOct4 in MEFs and mMyos are reprogrammed by oocyte factors (QPCR, Figure 8.2.3) to be lowly expressed at an oocyte-steady state (QPCR, Table

8.2.3 and 8.2.6; RNA-seq plus BrUTP pull-down, Figure 8.2.5.A and Table 8.2.7).

By taking all these into account, including effect of oocyte factors from Day 0 to Day 2 (QPCR) and expression level of genes at an oocyte-steady state (QPCR and RNA-seq plus BrUTP pull-down), mSall4, mUtf1 and mEsrrb in MEFs are mildly reprogrammed (QPCR, Table 8.2.5) by oocyte factors to low expression level at oocyte steady state (RNA-seq, Figure 8.2.6.B) and this means the resistance of these three genes is counteracting reprogramming by oocyte factors. The same comparison can also apply on mMyos and this shows that mUtf1, mKlf2, mSall4 and mEsrrb in mMyos are resistant to be reprogrammed by oocytes (Figure 8.2.6.C, Table 8.2.3, 8.2.5, 8.2.6 and 8.2.7).

If we now look at the results of overexpressing xklf2-HA, using the same RNA-seq plus BrUTP pull-down assay, these are shown in Figure 8.2.5.B and 8.2.6.D-F. For mESCs and MEF, only the lowest effect is seen. I should explain that in these figures, the values for expression of genes at an oocyte-steady state is set at 1.0 and the significant but very mild change from those values are for mSox2, mOct4, mKlf2 and mSall4 for both mESC and MEF donor cells (Figure 8.2.5.B, 8.2.6.D-E). Additionally, the mild change is also seen for mEsrrb in mESCs (Figure 8.2.6.D) and mUtf1 in MEFs (Figure 8.2.6.E). For mMyo cells, xklf2 has a very large effect on mSox2, mOct4, mKlf2 and mSall4 and a medium effect on mEsrrb (Figure 8.2.5.B and 8.2.6.F). Since xklf2-HA overexpression is most effective to activate reprogramming resistant genes from nil/low expression level by more fold

change, it means the resistance of mSox2, mOct4, mKlf2 and mSall4 in mMyos is stronger than resistance of these genes in mESCs and MEFs (Table 8.2.6.B). Therefore, overexpression of TFs in Oocyte-NT system can reveal the different levels of gene resistance between cell types.

In summary, certain genes, including mSox and mOct4, show a huge up-regulation by endogenous oocyte factors, with very little further up-regulation by overexpressed xklf2-HA when the genes are reprogrammed to an oocyte-steady state successfully by oocyte factors. A mild response to endogenous factors is seen for five other genes, and a weak or nil response by a further five genes. Again, the strong xklf2-HA overexpression effect is seen when genes resistant to activation by oocyte factors. This conclusion applies to the steady state level of RNA transcripts. When gene activation, as opposed to a steady state value is analysed, similar conclusions apply. Nuclei of mMyos show a similar response to those of MEF cells. It is also interesting that the pioneer factor mFox1a1 only enhances the expression of mSox2 and mOct4 but not other genes. Therefore, pioneer factor mFoxa1 cannot change the resistance of these other genes to reprogramming by oocytes.

Figure 8.2.1.A-B

Transcript assay: QPCR

Factor tested: Oocyte (maternal) and xklf2-HA overexpression

Gene response: A. mSox2 and mOct4

B. Other pluripotency genes

Cell type: MEF (sixiFM)

■ Expression level at Day 0 (carry-over of mouse transcripts)

■ Oocyte effect on mouse genes from Day 0 to Day 2 after Oocyte-NT

■ Effect of xklf2-HA overexpression on mouse genes at Day 2 after Oocyte-NT

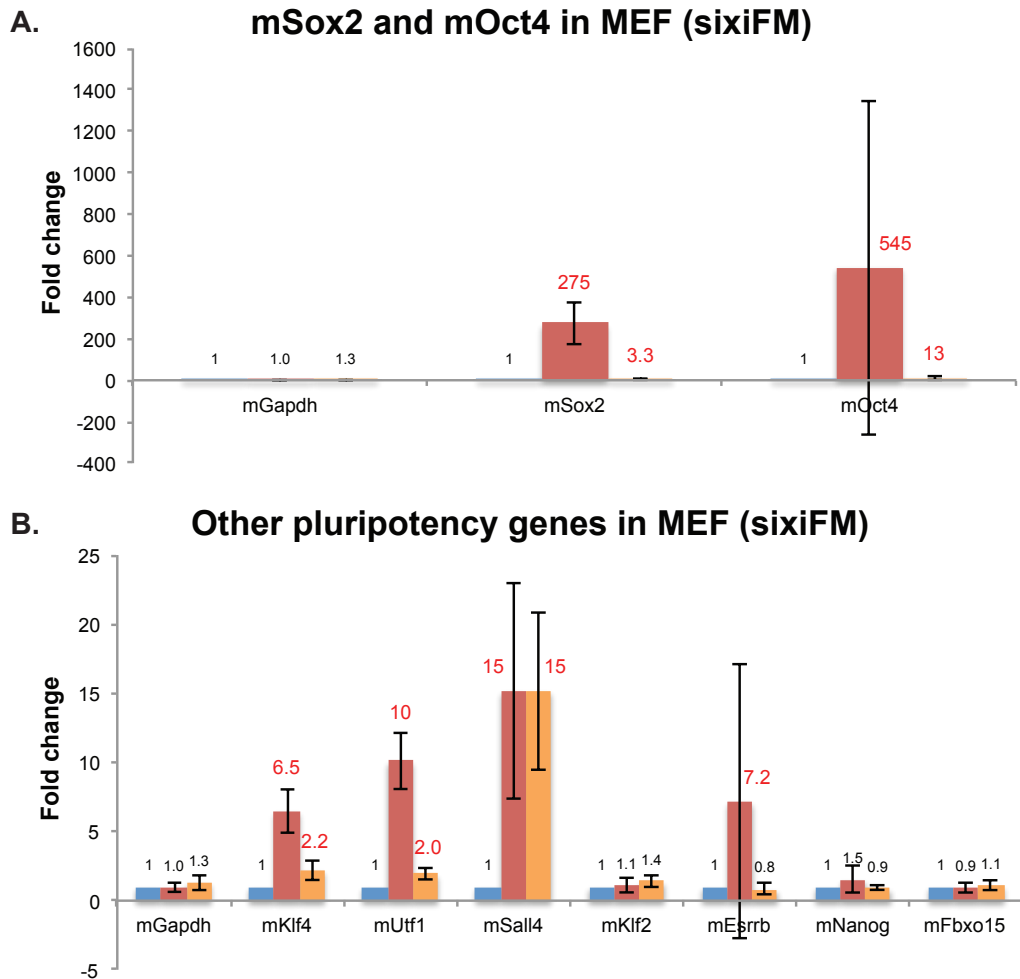


Figure 8.2.1.A-B In sixiFM MEFs, oocyte factors strongly up-regulate expression of mSox2 and mOct4 by up to 545-fold and mildly up-regulate expression of mKlf4, mUtf1, mSall4, mEsrrb from Day 0 to Day 2 after Oocyte-NT. xklf2-HA overexpression mildly up-regulates expression of mOct4 by 13-fold and weakly up-regulates expression of mSox2, mKlf4 and mUtf1 by up to 3.3-fold (sixiFM, n=3, Table 8.2.1).

Fold change more than 2 is shown in red.

Expression of mGapdh is not changed after Oocyte-NT and not affected by oocyte factors or xklf2-HA overexpression.

Figure 8.2.1.C-D

Transcript assay: QPCR

Factor tested: Oocyte (maternal) and xklf2-HA overexpression

Gene response: A. mSox2 and mOct4

B. Other pluripotency genes

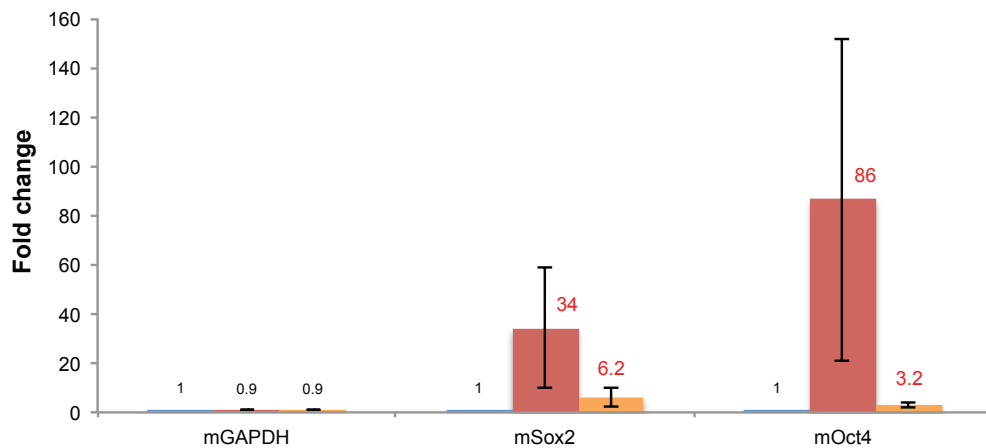
Cell type: MEF (TcR2)

■ Expression level at Day 0 (carry-over of mouse transcripts)

■ Oocyte effect on mouse genes from Day 0 to Day 2 after Oocyte-NT

■ Effect of xklf2-HA overexpression on mouse genes at Day 2 after Oocyte-NT

C. mSox2 and mOct4 in MEF (TcR2)



D. Other pluripotency genes in MEF (TcR2)

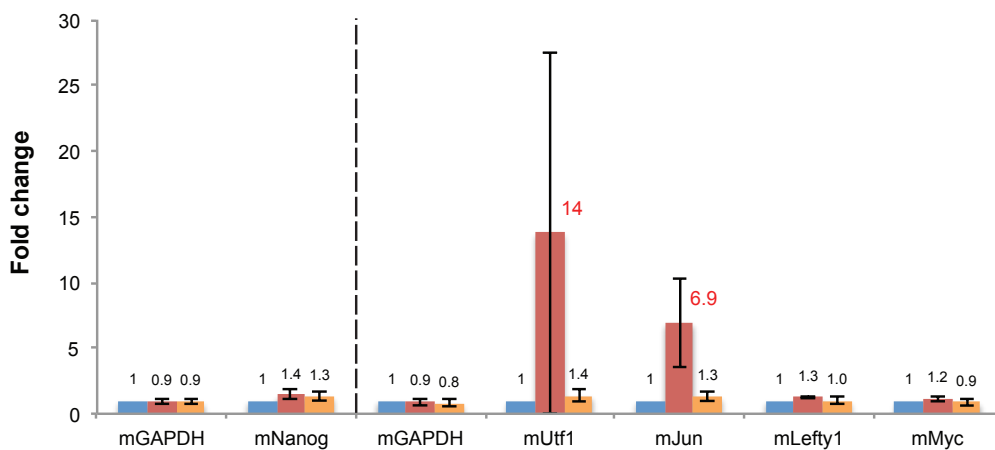


Figure 8.2.1.C-D In TcR2 MEFs, oocyte factors strongly up-regulate expression of mSox2 and mOct4 by up to 85-fold and mildly up-regulate expression of mUtf1 and mJun by up to 14-fold. xklf2-HA overexpression mildly up-regulates expression of mSox2 by 6.2-fold and weakly up-regulates expression of mOct4 by 3.2-fold (TcR2, n=2 or 3, Table 8.2.1).

Fold change more than 2 is shown in red.

Figure 8.2.2

Transcript assay: QPCR
Factor tested: Oocyte (maternal) and xklf2-HA overexpression
Gene response: A. mSox2 and mOct4
B. Other pluripotency genes
Cell type: MEF (sixiFM and TcR2)

- Expression level at Day 0 (carry-over of mouse transcripts)
- Oocyte effect on mouse genes from Day 0 to Day 3 after Oocyte-NT
- Effect of xklf2-HA overexpression on mouse genes at Day 3 after Oocyte-NT

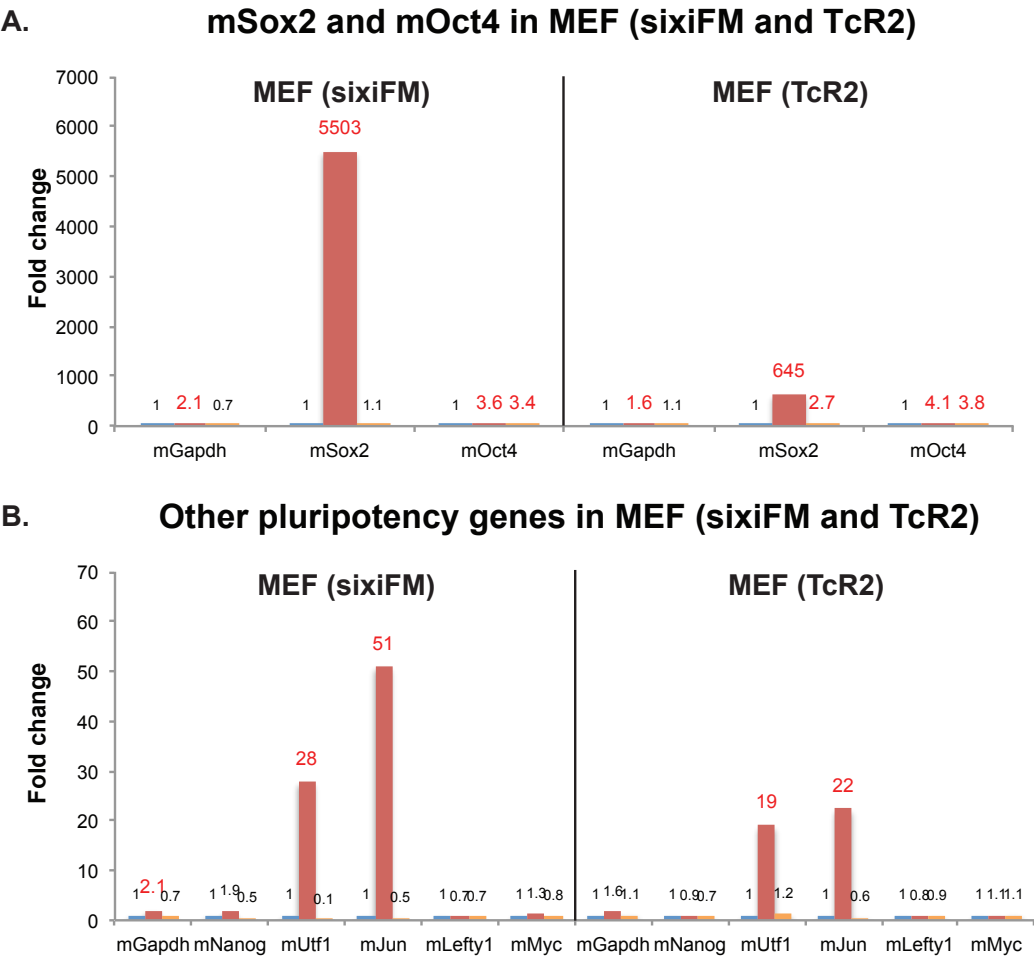


Figure 8.2.2 Pluripotency genes in different MEF cell lines, namely sixiFM and TcR2, respond to oocyte factors and xklf2-HA overexpression similarly (n=1, Table 8.2.2). Fold change more than 2 is shown in red.

Figure 8.2.3

Transcript assay: QPCR

Factor tested: Oocyte (maternal) and xklf2-HA overexpression

Gene response: A. mSox2 and mOct4

B. Other pluripotency genes

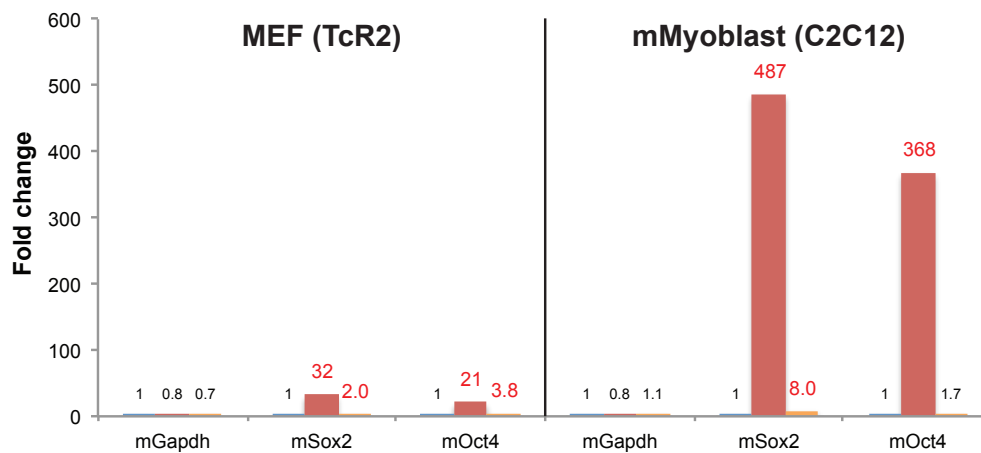
Cell type: MEF (TcR2) and mMyoblast (C2C12)

■ Expression level at Day 0 (carry-over of mouse transcripts)

■ Oocyte effect on mouse genes from Day 0 to Day 2 after Oocyte-NT

■ Effect of xklf2-HA overexpression on mouse genes at Day 2 after Oocyte-NT

A. mSox2 and mOct4 in MEF (TcR2) and mMyoblast (C2C12)



B. Other pluripotency genes in MEF (TcR2) and mMyoblast (C2C12)

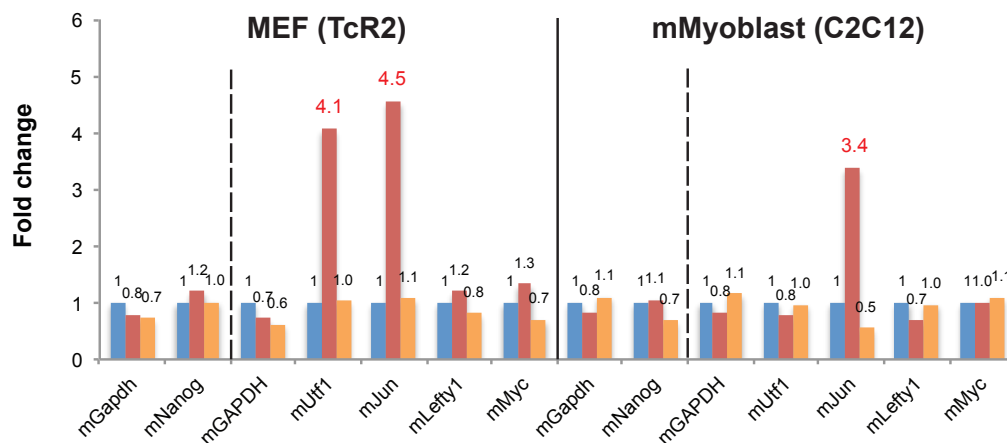


Figure 8.2.3 mSox2 and mOct4 in mMyoblast respond to oocyte factors stronger than MEF (n=1, Table 8.2.3).

Fold change more than 2 is shown in red. mUtf1 in mMyoblasts is not up-regulated by oocyte factors and mOct4 in mMyoblasts is not up-regulated by xklf2-HA overexpression due to SCNR resistance.

Figure 8.2.4

Transcript assay: QPCR

Factor tested: Oocyte (maternal) and overexpression of xklf2-HA and mFoxa1-HA

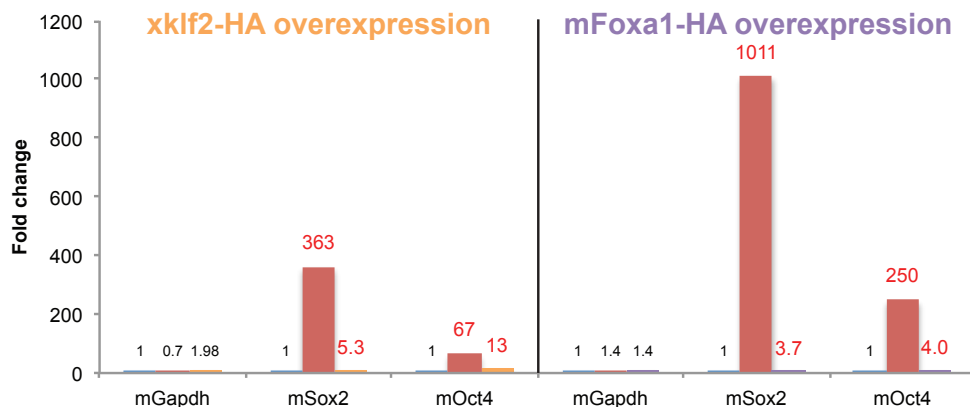
Gene response: A. mSox2 and mOct4

B. Other pluripotency genes

Cell type: MEF (sixiFM)

- Expression level at Day 0 (carry-over of mouse transcripts)
- Oocyte effects on mouse genes from Day 0 to 2 Day 2 after Oocyte-NT
- Effect of xklf2-HA overexpression on mouse genes at Day 2 after Oocyte-NT
- Effect of mFoxa1-HA overexpression on mouse genes at Day 2 after Oocyte-NT

A. mSox2 and mOct4 in MEF (sixiFM)



B. Other pluripotency genes in MEF (sixiFM)

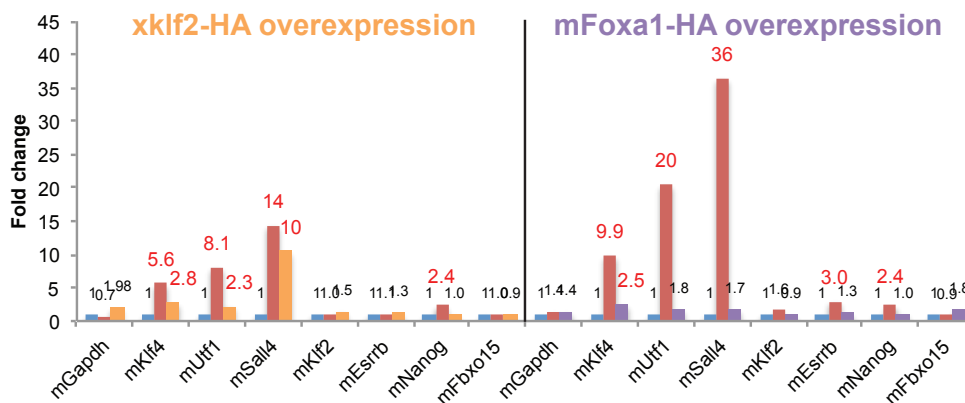


Figure 8.2.4 xklf2-HA overexpression can up-regulate more pluripotency genes than mFoxa1-HA overexpression (n=1, Table 8.2.4).

mEsrrb is not up-regulated by oocyte factors due to SCNR resistance. mNanog is only up-regulated by oocyte factors in this batch of oocytes but is not up-regulated in other oocyte batches. After normalization by mGapdh for cell number difference in each Oocyte-NT samples, mKlf4 and mUtf1 is un-affected by xklf2-HA and mFoxa1-HA overexpression.

Figure 8.2.5

Transcript assay: RNA-seq plus BrUTP pull-down

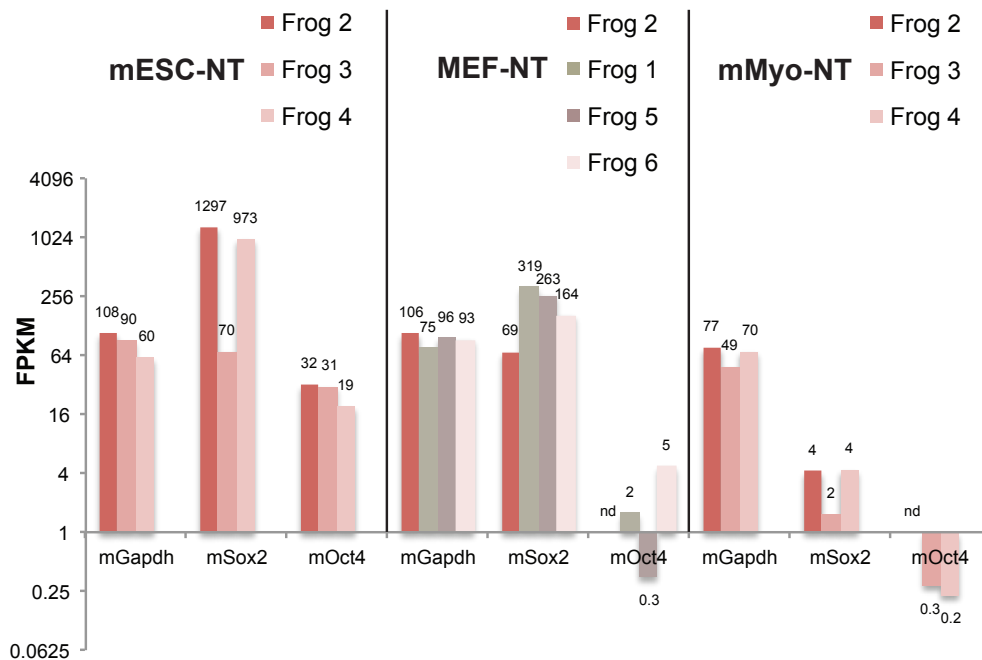
Factor tested: A. Oocyte (maternal)

B. *xklf2*-HA overexpression

Gene response: *mSox2* and *mOct4* at Day 2 after Oocyte-NT

Cell type: mESC, MEF (sixiFM), mMyo

A. Expression level of *mSox2* and *mOct4* at oocyte-steady state reprogrammed by maternal factors in mESC-NT, MEF-NT and mMyo-NT



B. Effect of *xklf2*-HA overexpression on *mSox2* and *mOct4* in mESC-NT, MEF-NT and mMyo-NT

■ Genes reprogrammed by maternal factors of *Xenopus* oocytes

■ Effect of *xklf2*-HA overexpression on reprogrammed genes

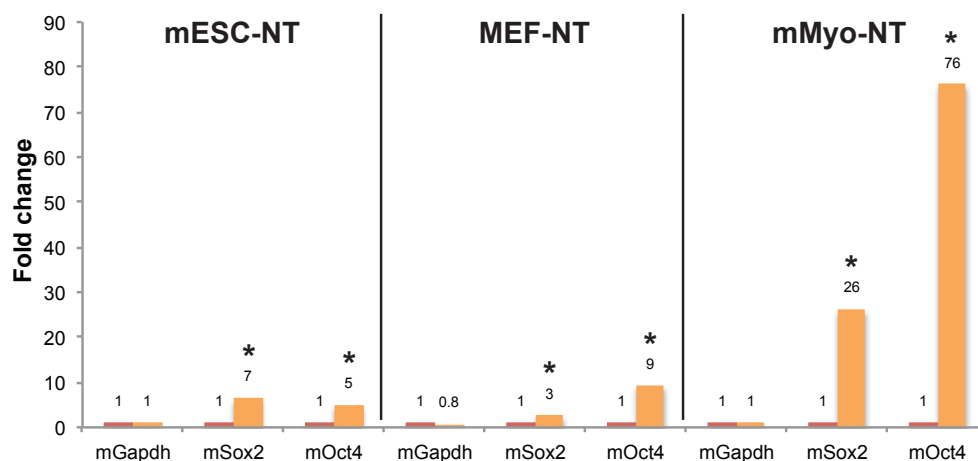


Figure 8.2.5 *mSox2* and *mOct4* in mMyo-NT is resistant to be activated by oocyte factors and hence respond strongly to *xklf2*-HA overexpression at Day 2 after Oocyte-NT.

nd, not detectable by RNA-seq, FPKM=0; *FDR<0.1

Figure 8.2.6.A-C

Transcript assay: RNA-seq plus BrUTP pull-down

Factor tested: Oocyte (maternal)

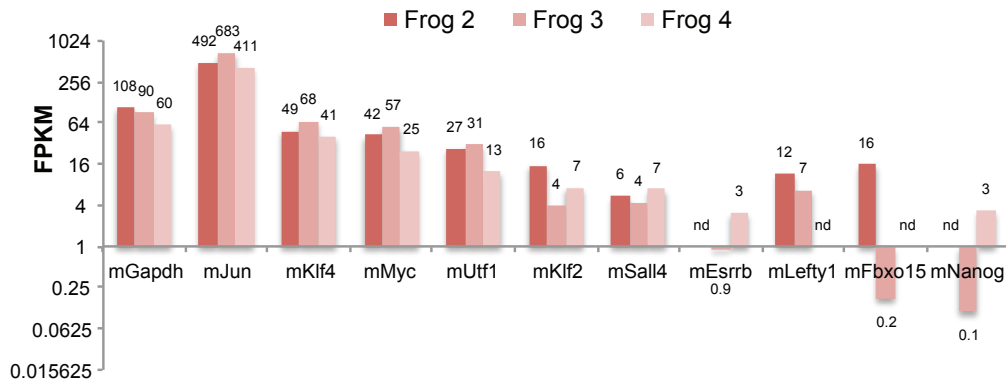
Gene response: Other pluripotency genes at Day 2 after Oocyte-NT

Cell type: A. mESC

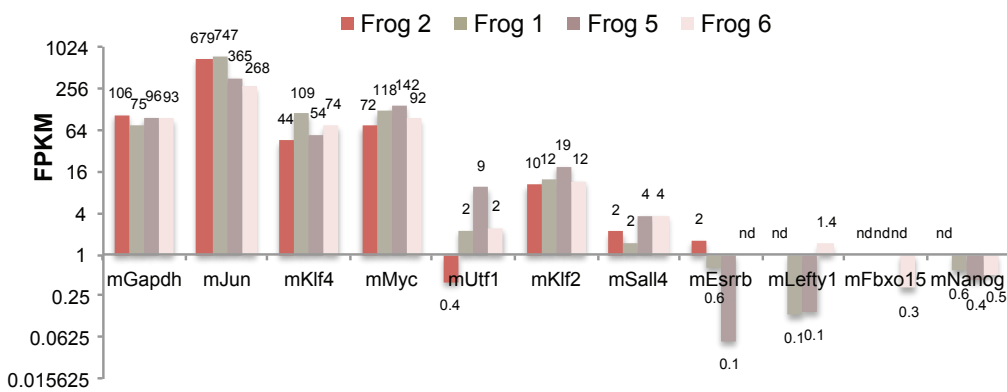
B. MEF (sixiFM)

C. mMyo

A. Expression level of other pluripotency genes at oocyte-steady state reprogrammed by oocyte factors in mESC-NT



B. Expression level of other pluripotency genes at oocyte-steady state reprogrammed by oocyte factors in MEF-NT



C. Expression level of other pluripotency genes at oocyte-steady state reprogrammed by oocyte factors in mMyo-NT

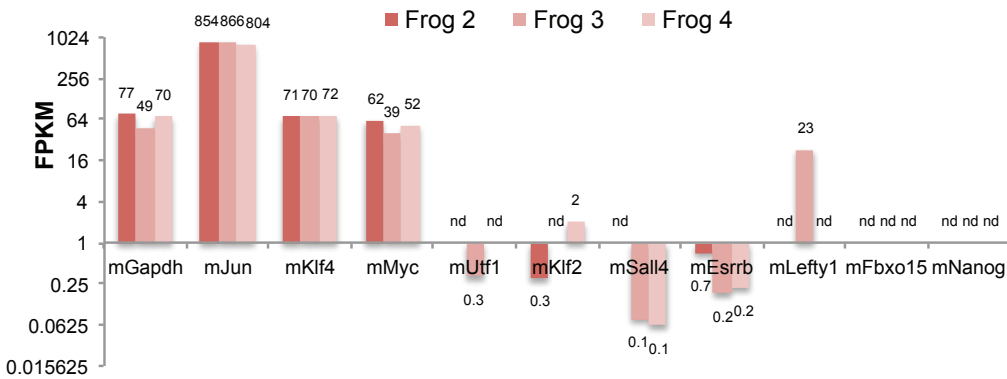


Figure 8.2.6.D-F

Transcript assay: RNA-seq plus BrUTP pull-down

Factor tested: xklf2-HA overexpression

Gene response: Other pluripotency genes at Day 2 after Oocyte-NT

Cell type: D. mESC

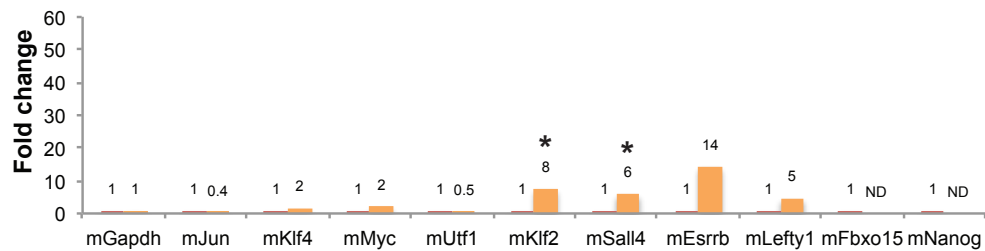
E. MEF (sixiFM)

F. mMyo

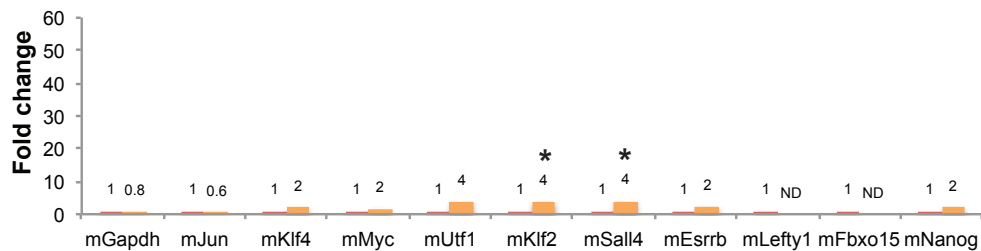
■ Genes reprogrammed by maternal factors of *Xenopus* oocytes

■ Effect of xklf2-HA overexpression on reprogrammed genes

D. Effect of xklf2-HA overexpression at oocyte-steady state on other pluripotency genes in mESC-NT



E. Effect of xklf2-HA overexpression on at oocyte-steady state other pluripotency genes in MEF-NT



F. Effect of xklf2-HA overexpression on at oocyte-steady state other pluripotency genes in mMyo-NT

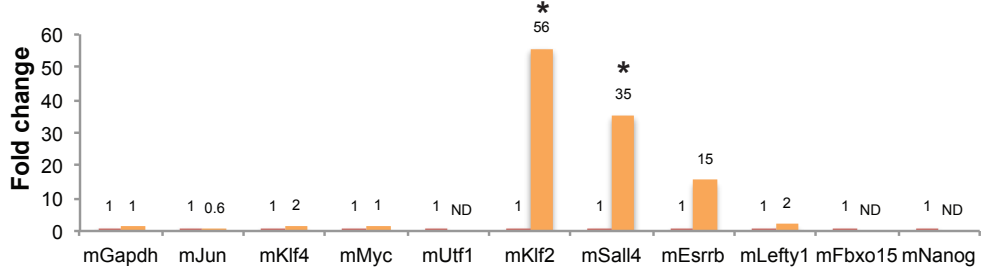


Figure 8.2.6 The highest up-regulation of mKlf2 and mSall4 in mMyo-NT induced by xklf2-HA overexpression indicates mKlf2 and mSall4 in mMyo-NT have strongest SCNR resistance than the same genes in mESC-NT and MEF-NT.

(A-C) Most pluripotency genes are expressed highly in mESC-NT while some pluripotency genes in mMyo-NT fail to be activated by oocyte factors.

nd, not detectable by RNA-seq, FPKM=0

(D-E) mKlf2 and mSall4 in mMyo-NT respond most strongly to xklf2-HA overexpression due to the strongest SCNR resistance of them than those in mESC-NT and MEF-NT.

ND, not determined; *FDR<0.1

Table 8.2.1

Overexpressed TF	Frog no.	Gene	Gene expression (Ct) and culture period			Fold change by		Cell type
			Oocyte factors		Oocyte factors + TF overexpression	Oocyte factors (Day 0 to Day 2)	TF overexpression (Day 2)	
			Day 0	Day 2				
xklf2-HA	1	mGapdh	25	25	25	0.9	1.1	MEF (sixiFM)
		mSox2	34	27	25	165	3.0	
		mOct4	40*	34	29	91	24	
		mKlf4	30	28	26	5.5	2.5	
		mUtf1	34	30	29	12	2.1	
		mSall4	34	31	26	7.8	22	
		mKlf2	27	27	26	0.7	1.8	
		mEsrrb	32	31	31	1.9	0.8	
		mNanog	26	27	27	0.5	0.8	
	mFbxo15	31	32	32	0.5	1.6		
	2	mGapdh	27	28	27	0.7	2.0	
		mSox2	37	28	26	363	5.3	
		mOct4	43*	37	33	67	13	
		mKlf4	32	30	28	5.6	2.8	
		mUtf1	34	31	30	8.1	2.3	
		mSall4	37	33	30	14	10	
		mKlf2	30	30	29	1.0	1.5	
		mEsrrb	34	34	34	1.1	1.3	
		mNanog	33	32	32	2.4	1.0	
	mFbxo15	35	35	35	1.0	0.9		
	3	mGapdh	24	24	24	1.4	0.9	
		mSox2	33	24	24	296	1.8	
		mOct4	39*	29	28	1476	1.3	
		mKlf4	29	26	26	8.3	1.4	
		mUtf1	32	29	28	10	1.5	
		mSall4	34	29	25	23	14	
		mKlf2	27	26	26	1.7	1.0	
		mEsrrb	32	28	29	19	0.5	
		mNanog	26	26	26	1.7	0.9	
	mFbxo15	32	32	32	1.3	0.8		
	4	mGAPDH	21	21	21	1.1	1.0	MEF (TcR2)
		mSox2	29	25	22	11	9.6	
		mOct4	36	30	29	86	1.8	
		mNanog	30	29	28	1.8	1.3	
		mUtf1	28	24	23	24	1.7	
		mJun	28	25	24	9.3	1.5	
		mLefty1	24	24	24	1.3	1.2	
		mMyc	17	17	17	1.0	1.1	
	5	mGAPDH	21	22	22	0.8	0.7	
		mSox2	29	24	23	32	2.02	
		mOct4	36	32	30	21	3.8	
		mNanog	27	27	27	1.2	1.0	
mGAPDH		25	26	26	0.7	0.6		
mUtf1		32	30	30	4.1	1.0		
mJun		31	29	29	4.5	1.1		
mLefty1		28	28	28	1.2	0.8		
mMyc	22	21	22	1.3	0.7			
6	mGAPDH	22	22	22	0.9	1.0		
	mSox2	30	24	22	59	7.1		
	mOct4	38	31	29	152	4.1		
	mNanog	26	26	25	1.2	1.7		

Table 8.2.1 QPCR analysis shows the effects of maternal factors from Day 0 to Day 2 and xklf2-HA overexpression at Day 2 after Oocyte-NT on pluripotency genes in MEF (sixiFM, n=3) and MEF (TcR2, n=2 or 3).

Fold change more than 2 is shown in red. Fold change more than 4 has background in pink.

[#] For each sample, Ct value for mOct4 in MEF (sixiFM) at 0 hour (in bold with underline) is assigned as Ct value for mGapdh plus 15, referring to the Ct value for mOct4 in MEF (TcR2) while the average Ct difference between mOct4 and mGapdh in MEF (TcR2) is 15.

Table 8.2.2

Overexpressed TF	Frog no.	Gene	Gene expression (Ct) and culture period			Fold change by		Cell type
			Oocyte factors		Oocyte factors + TF overexpression	Oocyte factors (Day 0 to Day 3)	TF overexpression (Day 3)	
			Day 0	Day 3	Day 3			
xklf2-HA	7	mGapdh	21	20	20	2.1	0.7	MEF (sixiFM)
		mSox2	35	22	22	5503	1.1	
		mOct4	36*	34	32	3.6	3.4	
		mNanog	29	28	29	1.9	0.5	
		mUtf1	29	24	25	28	0.1	
		mJun	27	22	22	51	0.5	
		mLefty1	29	29	30	0.7	0.7	
		mMyc	17	17	17	1.3	0.8	MEF (TcR2)
		mGapdh	20	20	19	1.6	1.1	
		mSox2	31	22	20	645	2.7	
		mOct4	35*	33	31	4.1	3.8	
		mNanog	28	29	29	0.9	0.7	
		mUtf1	29	25	24	19	1.2	
		mJun	26	22	22	22	0.6	
		mLefty1	28	29	29	0.8	0.9	
		mMyc	17	17	17	1.1	1.1	

Table 8.2.2 QPCR analysis shows the effects of oocyte factors from Day 0 to Day 3 after Oocyte-NT and xklf2-HA overexpression at Day 3 after Oocyte-NT on pluripotency genes in MEF (sixiFM, n=1) and MEF (TcR2, n=1).

Fold change more than 2 is shown in red. Fold change more than 4 has background in pink.

[#] Ct value for mOct4 in MEF (sixiFM and TcR2) at 0 hour (in bold with underline) is assigned as Ct for mGapdh plus 15, referring to the average Ct difference between mOct4 and mGapdh in MEF (TcR2) in Table 8.2.1.

Table 8.2.3

Overexpressed TF	Frog no.	Gene	Gene expression (Ct) and culture period			Fold change by		Cell type
			Oocyte factors		Oocyte factors + TF overexpression	Oocyte factors (Day 0 to Day 2)	TF overexpression (Day 2)	
			Day 0	Day 2				
xklf2-HA	5	mGapdh	21	22	22	0.8	0.7	MEF (TcR2)
		mSox2	29	24	23	32	2.0	
		mOct4	36	32	30	21	3.8	
		mNanog	27	27	27	1.2	1.0	mMyoblas (C2C12)
		mGapdh	21	21	21	0.8	1.1	
		mSox2	39	30	27	487	8.0	
		mOct4	39	30	29	368	1.7	
		mNanog	26	26	27	1.1	0.7	MEF (TcR2)
		mGAPDH	25	26	26	0.7	0.6	
		mUtf1	32	30	30	4.1	1.0	
		mJun	31	29	29	4.5	1.1	
		mLefty1	28	28	28	1.2	0.8	
		mMyc	22	21	22	1.3	0.7	mMyoblas (C2C12)
		mGAPDH	25	25	25	0.8	1.1	
		mUtf1	32	32	32	0.8	1.0	
		mJun	31	29	30	3.4	0.5	
		mLefty1	28	29	29	0.7	1.0	
		mMyc	22	22	21	1.0	1.1	

Table 8.2.3 QPCR analysis shows the effects of oocyte factors from Day 0 to Day 2 after Oocyte-NT and xklf2-HA overexpression at Day 2 after Oocyte-NT on pluripotency genes in MEF (TcR2, n=1) and mMyoblast (C2C12, n=1).

Fold change more than 2 is shown in red. Fold change more than 4 has background in pink.

Table 8.2.4

Overexpressed TF	Frog no.	Gene	Gene expression (Ct) and culture period			Fold change by		Cell type
			Oocyte factors		Oocyte factors + TF overexpression	Oocyte factors (Day 0 to Day 2)	TF overexpression (Day 2)	
			Day 0	Day 2				
xklf2-HA	2	mGapdh	27	28	27	0.7	1.98	MEF (sixiFM)
		mSox2	37	28	26	363	5.3	
		mOct4	43	37	33	67	13	
		mKlf4	32	30	28	5.6	2.8	
		mUtf1	34	31	30	8.1	2.3	
		mSall4	37	33	30	14	10	
		mKlf2	30	30	29	1.0	1.5	
		mEsrrb	34	34	34	1.1	1.3	
		mNanog	33	32	32	2.4	1.0	
		mFbxo15	35	35	35	1.0	0.9	
mFoxa1-HA	mGapdh	27	27	26	1.4	1.4		
	mSox2	37	27	25	1011	3.7		
	mOct4	43	35	33	250	4.0		
	mKlf4	32	29	28	9.9	2.5		
	mUtf1	34	30	29	20	1.8		
	mSall4	37	32	31	36	1.7		
	mKlf2	30	29	29	1.6	0.9		
	mEsrrb	34	33	32	3.0	1.3		
	mNanog	33	32	32	2.4	1.0		
	mFbxo15	35	35	34	0.9	1.8		

Table 8.2.4 QPCR analysis shows the effects of oocyte factors from Day 0 to 2 Day 2 after Oocyte-NT and TF overexpression (xklf2-HA and mFoxa1-HA) at Day 2 after Oocyte-NT on pluripotency genes in MEF (sixiFM, n=1).

Fold change more than 2 is shown in red. Fold change more than 4 has background in pink.

Ct value for mOct4 in MEF (sixiFM and TcR2) at 0 hour (in bold with underline) is assigned as Ct for mGapdh plus 15, referring to the average Ct difference between mOct4 and mGapdh in MEF (TcR2) in Table 8.2.1.

Table 8.2.5**Transcript assay: QPCR****Factor tested: Oocyte (maternal) and xklf2-HA overexpression****Gene response: level of reponse for all pluripotency genes****Cell type: MEF (sixiFM and TcR2)**

QPCR (From Table 8.2.1)					
Response to Oocyte factors	Cell type	Gene	Frog no.	Average fold change	
				Oocyte factors (Day 0 to Day 2)	xklf2-HA overexpression (Day 2)
Nil	MEF (sixiFM and TcR2)	mGapdh	1-6	0.9	1.1
Strong	MEF (sixiFM and TcR2)	mSox2	1-6	154	4.8
	MEF (TcR2)	mOct4	4-6	86	8.2
Mild	MEF (sixiFM)	mSall4	1-3	15	15
	MEF (sixiFM and TcR2)	mUtf1	1-5	12	1.7
	MEF (sixiFM)	mEsrrb	1-3	7.2	0.8
	MEF (TcR2)	mJun	4-5	6.9	1.3
	MEF (sixiFM)	mKlf4	1-3	6.5	2.2
Weak/Nil	MEF (sixiFM and TcR2)	mNanog	1-6	1.5	1.1
	MEF (TcR2)	mLefty1	4-5	1.3	1.0
	MEF (TcR2)	mMyc	4-5	1.2	0.9
	MEF (sixiFM)	mKlf2	1-3	1.1	1.4
	MEF (sixiFM)	mFbxo15	1-3	0.9	1.1

Table 8.2.5 Summary of QPCR analysis shows that, in MEF (sixiFM and TcR2), mSox2 and mOct4 are strongly up-regulated by oocyte factors with average fold change from 86 to 154 and 5 other tested pluripotency genes are mildly up-regulated. Comparably, 3 tested pluripotency genes mildly respond to xklf2-HA overexpression with average fold change from 4.8 to 15 and mKlf4 weakly responds to xklf2-HA overexpression with average fold change of 2.2.

Table 8.2.6

Transcript assay: QPCR

Factor tested: A. Oocyte (maternal)

B. xklf2-HA overexpression

Gene response: level of response for all pluripotency genes at Day 2 after Oocyte-NT

Cell type: MEF (sixiFM and TcR2), mMyo

A. Effect of oocyte factors at Day 2 after Oocyte-NT

Cell type (Tables)	Reprogrammed expression level of pluripotency genes and mGapdh					
	High ($Ct_{mGapdh} - Ct_{Gene} \leq 3$)		Medium $3 < (Ct_{mGapdh} - Ct_{Gene}) \leq 6$		Low/nil ($Ct_{mGapdh} - Ct_{Gene} > 6$)	
	No. of genes	Gene	No. of genes	Gene	No. of genes	Gene
MEF (8.2.1)	6/13	mGapdh mSox2 mKlf4 mMyc mKlf2 mLefty1	5/13	mJun mUtf1 mSall4 mEsrrb mNanog	2/13	mOct4 mFbxo15
mMyo (8.2.3)	2/8	mGapdh mMyc	3/8	mJun mLefty1 mNanog	3/8	mSox2 mOct4 mUtf1

B. Effect of xklf2-HA overexpression at Day 2 after Oocyte-NT

Cell type (Figures)	Effect of xklf2-HA overexpression				Effect of mFxa1-HA overexpression			
	Up-regulated (Fold change ≥ 2)		Un-affected (Fold change < 2)		Up-regulated (Fold change ≥ 2)		Un-affected (Fold change < 2)	
	No. of genes	Gene	No. of genes	Gene	No. of genes	Gene	No. of genes	Gene
MEF (8.2.1)	3/13	mSox2 mOct4 mSall4	10/13	mGapdh mJun mKlf4 mMyc mUtf1 mKlf2 mEsrrb mLefty1 mFbxo15 mNanog	2/9	mSox2 mOct4	7/9	mGapdh mUtf1 mKlf2 mSall4 mEsrrb mFbxo15 mNanog
mMyo (8.2.3)	1/8	mSox2	7/8	mGapdh mOct4 mNanog mUtf1 mJun mLefty1 mMyc				

Table 8.2.6 QPCR analysis measures the expression level of pluripotency genes in MEFs and mMyos reprogrammed by oocytes factors at Day 2 after Oocyte-NT (A) and response of these genes to xklf2-HA overexpression at Day 2 after Oocyte-NT (B).

Table 8.2.7

Transcript assay: RNA-seq plus BrUTP pull-down

Factor tested: A. Oocyte (maternal)

B. xklf2-HA overexpression

Gene response: level of response for all pluripotency genes at Day 2 after Oocyte-NT

Cell type: mESC, MEF (sixiFM), mMyo

A. Effect of oocyte factors at Day 2 after Oocyte-NT

Cell type (Figures)	Reprogrammed expression level of pluripotency genes and mGapdh					
	High (FPKM>4)		Medium (1≤FPKM≤4)		Low/nil (FPKM<1)	
	No. of genes	Gene	No. of genes	Gene	No. of genes	Gene
mESC (8.2.5.A and 8.2.6.A)	7/13	mGapdh mSox2 mOct4 mJun mKlf4 mMyc mUtf1	2/13	mKlf2 mSall4	4/13	mEsrrb mLefty1 mFbxo15 mNanog
MEF (8.2.5.A and 8.2.6.B)	6/13	mGapdh mSox2 mJun mKlf4 mMyc mKlf2	1/13	mSall4	6/13	mOct4 mUtf1 mEsrrb mLefty1 mFbxo15 mNanog
mMyo (8.2.5.A and 8.2.6.C)	4/13	mGapdh mJun mKlf4 mMyc	1/13	mSox2	8/13	mOct4 mUtf1 mKlf2 mSall4 mEsrrb mLefty1 mFbxo15 mNanog

B. Effect of xklf2-HA overexpression at Day 2 after Oocyte-NT

Cell type (Figures)	Resistance of pluripotency genes (Fold change increased by xklf2-HA overexpression)			
	mSox2	mOct4	mKlf2	mSall4
mESC (8.2.5.B and 8.2.6.D)	Middle (7)	Weak (5)	Middle (8)	Middle (6)
MEF (8.2.5.B and 8.2.6.E)	Weak (3)	Middle (9)	Weak (4)	Weak (4)
mMyo (8.2.5.B and 8.2.6.F)	Strong (26)	Strong (76)	Strong(56)	Strong (35)

Table 8.2.7 RNA-seq plus BrUTP pull-down measures the relative amount of newly synthesized transcripts (FPKM) at an oocyte steady state at Day 2 after Oocyte-NT. SCNR resistance of pluripotency genes can be judged by response of different cell types to xklf2-HA overexpression at Day 2 after Oocyte-NT.

8.3 Cell-hybrid experiments

Cell fusion or cell hybrids were first made by Okada, Y¹⁵⁶⁻¹⁵⁸ and gained wide interest after the results of Harris, H^{159,160} in the 1960s. There are several designs of experiments. In one, two whole single-cells are fused such that there are, initially, two nuclei enclosed in the mixed cytoplasm of two cells. These hybrids can be grown with much cell division and much selection. They are not particularly useful for analyzing early events that affect gene expression by these hybrids.

Another design of experiment yields heterokaryons. These, again, are initially two nuclei in the fused cytoplasm of two cells, but these hybrid cells are not grown, but are cultured for a few days. These are the most useful experiments for comparison with nuclear transfer to amphibian oocytes. In some cases, a cell can be fused with the enucleated cytoplasm of another cell, called a cytoplast. Also, the ratio of number of nuclei to number of cytoplasts can be varied. Also, the amount of cytoplasm relative to nucleus can be varied by using the nuclei of red blood cells, which are very small, fused to cultured cells which are usually large. The two laboratories which have done the greatest amount of work with heterokaryons are those of Fisher (London) and Blau (Stanford). Commonly, these experiments involved combining cells of two different species so that genes in one species can be seen to be activated by the cytoplasm of another.

An early impressive experiment was that of Blau. This involved the fusion of a human hepatocyte with a mouse muscle cell^{42,44}. It was found that, in this

heterokaryon, a human muscle antigen was expressed due to the inclusion of cytoplasm from a mouse muscle cell. In this case, the amount of mouse cytoplasm far exceeded that of the human hepatocyte cytoplasm because mouse cells are multinucleate. An advantage of this design of experiments is that the hybrid cells are not allowed to replicate DNA or divide (use of inhibitors) so that there is no loss of chromosomes during the incubation period. Also, mouse and human muscle genes can be distinguished by antibodies.

Some of the points which emerge from this heterokaryon work are

1. Gene activation is much influenced by an excess of cytoplasm from one cell partner to another.
2. The two kinds of nuclei do not fuse and therefore can be readily distinguished, thereby giving attention to only those heterokaryons which contain one original cell of each type.
3. Cell-type specific factors from one kind of cell operate functionally on another kind of cell, even if both are from unrelated animal species.
4. The new gene activation takes place within a few days.

There has been much discussion in the literature on the question of whether DNA synthesis and/or cell division are required for the transactivation of the gene of one species by the cytoplasm of another. In the experiments of Blau, it is clear that DNA synthesis and cell division are not required^{42,44}. Conversely, in the experiments of Fisher, it is maintained that DNA synthesis and cell division are absolutely necessary¹⁶¹. For comparison, it is very clear

that, in nuclear injection to oocytes, no DNA synthesis or cell division is involved.

In the experiments of Fisher, it is argued that the discontinuation of cell-type specific gene expression is involved in any reprogramming^{162,163}. It is also argued that the gene Oct4 has to be expressed in the reprogramming process for other genes to be activated¹⁶². The extent and frequency of gene activation is small unless the heterokaryons can be isolated and looked at individually¹⁶².

Chapter 9 References

- 1 Buganim, Y., Faddah, D. A. & Jaenisch, R. Mechanisms and models of somatic cell reprogramming. *Nature reviews. Genetics* **14**, 427-439, doi:10.1038/nrg3473 (2013).
- 2 Vierbuchen, T. & Wernig, M. Molecular roadblocks for cellular reprogramming. *Molecular cell* **47**, 827-838, doi:10.1016/j.molcel.2012.09.008 (2012).
- 3 Shi, Y., Inoue, H., Wu, J. C. & Yamanaka, S. Induced pluripotent stem cell technology: a decade of progress. *Nat Rev Drug Discov* **16**, 115-130, doi:10.1038/nrd.2016.245 (2017).
- 4 Jopling, C., Boue, S. & Izpisua Belmonte, J. C. Dedifferentiation, transdifferentiation and reprogramming: three routes to regeneration. *Nature reviews. Molecular cell biology* **12**, 79-89, doi:10.1038/nrm3043 (2011).
- 5 Matoba, S. & Zhang, Y. Somatic Cell Nuclear Transfer Reprogramming: Mechanisms and Applications. *Cell stem cell* **23**, 471-485, doi:10.1016/j.stem.2018.06.018 (2018).
- 6 Lu, F. & Zhang, Y. Cell totipotency: molecular features, induction, and maintenance. *Natl Sci Rev* **2**, 217-225, doi:10.1093/nsr/nwv009 (2015).
- 7 Hochedlinger, K. & Jaenisch, R. Monoclonal mice generated by nuclear transfer from mature B and T donor cells. *Nature* **415**, 1035-1038, doi:10.1038/nature718 (2002).
- 8 Eggan, K. *et al.* Mice cloned from olfactory sensory neurons. *Nature* **428**, 44-49, doi:10.1038/nature02375 (2004).
- 9 Gurdon, J. B. & Wilmot, I. Nuclear transfer to eggs and oocytes. *Cold Spring Harb Perspect Biol* **3**, doi:10.1101/cshperspect.a002659 (2011).
- 10 Briggs, R. & King, T. J. Transplantation of Living Nuclei from Blastula Cells into Enucleated Frogs Eggs. *Proceedings of the National Academy of Sciences of the United States of America* **38**, 455-463, doi:DOI 10.1073/pnas.38.5.455 (1952).
- 11 King, T. J. & Briggs, R. Changes in the Nuclei of Differentiating Gastrula Cells, as Demonstrated by Nuclear Transplantation. *Proceedings of the National Academy of Sciences of the United States of America* **41**, 321-325 (1955).
- 12 Gurdon, J. B., Elsdale, T. R. & Fischberg, M. Sexually mature individuals of *Xenopus laevis* from the transplantation of single somatic nuclei. *Nature* **182**, 64-65 (1958).
- 13 Gurdon, J. B. The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. *J Embryol Exp Morphol* **10**, 622-640 (1962).
- 14 Gurdon, J. B. Adult frogs derived from the nuclei of single somatic cells. *Developmental biology* **4**, 256-273 (1962).
- 15 Gurdon, J. B. & Uehlinger, V. "Fertile" intestine nuclei. *Nature* **210**, 1240-1241 (1966).
- 16 Wilmot, I., Schnieke, A. E., McWhir, J., Kind, A. J. & Campbell, K. H. S. Viable offspring derived from fetal and adult mammalian cells. *Nature* **385**, 810-813, doi:DOI 10.1038/385810a0 (1997).

- 17 Liu, Z. *et al.* Cloning of Macaque Monkeys by Somatic Cell Nuclear Transfer. *Cell* **174**, 245, doi:10.1016/j.cell.2018.01.036 (2018).
- 18 Cibelli, J. B. & Gurdon, J. B. Custom-Made Oocytes to Clone Non-human Primates. *Cell* **172**, 647-649, doi:10.1016/j.cell.2018.01.030 (2018).
- 19 Alberio, R., Johnson, A. D., Stick, R. & Campbell, K. H. Differential nuclear remodeling of mammalian somatic cells by *Xenopus laevis* oocyte and egg cytoplasm. *Experimental cell research* **307**, 131-141, doi:10.1016/j.yexcr.2005.02.028 (2005).
- 20 Bian, Y., Alberio, R., Allegrucci, C., Campbell, K. H. & Johnson, A. D. Epigenetic marks in somatic chromatin are remodelled to resemble pluripotent nuclei by amphibian oocyte extracts. *Epigenetics* **4**, 194-202 (2009).
- 21 Miyamoto, K. *et al.* Cell-free extracts from mammalian oocytes partially induce nuclear reprogramming in somatic cells. *Biology of reproduction* **80**, 935-943, doi:10.1095/biolreprod.108.073676 (2009).
- 22 Bui, H. T. *et al.* Epigenetic reprogramming in somatic cells induced by extract from germinal vesicle stage pig oocytes. *Development* **139**, 4330-4340, doi:10.1242/dev.086116 (2012).
- 23 Yamanaka, S. & Blau, H. M. Nuclear reprogramming to a pluripotent state by three approaches. *Nature* **465**, 704-712, doi:10.1038/nature09229 (2010).
- 24 Martello, G. & Smith, A. The nature of embryonic stem cells. *Annu Rev Cell Dev Biol* **30**, 647-675, doi:10.1146/annurev-cellbio-100913-013116 (2014).
- 25 Boskovic, A. *et al.* Higher chromatin mobility supports totipotency and precedes pluripotency in vivo. *Genes & development* **28**, 1042-1047, doi:10.1101/gad.238881.114 (2014).
- 26 Hanna, J. H., Saha, K. & Jaenisch, R. Pluripotency and cellular reprogramming: facts, hypotheses, unresolved issues. *Cell* **143**, 508-525, doi:10.1016/j.cell.2010.10.008 (2010).
- 27 Nichols, J. & Smith, A. Naive and primed pluripotent states. *Cell stem cell* **4**, 487-492, doi:10.1016/j.stem.2009.05.015 (2009).
- 28 Gafni, O. *et al.* Derivation of novel human ground state naive pluripotent stem cells. *Nature* **504**, 282-+, doi:10.1038/nature12745 (2013).
- 29 Takashima, Y. *et al.* Resetting transcription factor control circuitry toward ground-state pluripotency in human. *Cell* **158**, 1254-1269, doi:10.1016/j.cell.2014.08.029 (2014).
- 30 Warrier, S. *et al.* Direct comparison of distinct naive pluripotent states in human embryonic stem cells. *Nature communications* **8**, 15055, doi:10.1038/ncomms15055 (2017).
- 31 Stadtfeld, M. & Hochedlinger, K. Induced pluripotency: history, mechanisms, and applications. *Genes & development* **24**, 2239-2263, doi:10.1101/gad.1963910 (2010).
- 32 Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663-676, doi:10.1016/j.cell.2006.07.024 (2006).

- 33 Okita, K., Ichisaka, T. & Yamanaka, S. Generation of germline-competent induced pluripotent stem cells. *Nature* **448**, 313-317, doi:10.1038/nature05934 (2007).
- 34 Hanna, J. *et al.* Direct reprogramming of terminally differentiated mature B lymphocytes to pluripotency. *Cell* **133**, 250-264, doi:10.1016/j.cell.2008.03.028 (2008).
- 35 Choi, J. *et al.* A comparison of genetically matched cell lines reveals the equivalence of human iPSCs and ESCs. *Nature biotechnology* **33**, 1173-1181, doi:10.1038/nbt.3388 (2015).
- 36 Do, J. T., Han, D. W. & Scholer, H. R. Reprogramming somatic gene activity by fusion with pluripotent cells. *Stem Cell Rev* **2**, 257-264, doi:10.1007/BF02698052 (2006).
- 37 Do, J. T. & Scholer, H. R. Nuclei of embryonic stem cells reprogram somatic cells. *Stem cells* **22**, 941-949, doi:10.1634/stemcells.22-6-941 (2004).
- 38 Silva, J., Chambers, I., Pollard, S. & Smith, A. Nanog promotes transfer of pluripotency after cell fusion. *Nature* **441**, 997-1001, doi:10.1038/nature04914 (2006).
- 39 Freberg, C. T., Dahl, J. A., Timoskainen, S. & Collas, P. Epigenetic reprogramming of OCT4 and NANOG regulatory regions by embryonal carcinoma cell extract. *Molecular biology of the cell* **18**, 1543-1553, doi:10.1091/mbc.E07-01-0029 (2007).
- 40 Foshay, K. M. *et al.* Embryonic stem cells induce pluripotency in somatic cell fusion through biphasic reprogramming. *Molecular cell* **46**, 159-170, doi:10.1016/j.molcel.2012.02.013 (2012).
- 41 Aloia, L., McKie, M. A. & Huch, M. Cellular plasticity in the adult liver and stomach. *The Journal of physiology* **594**, 4815-4825, doi:10.1113/JP271769 (2016).
- 42 Blau, H. M., Chiu, C. P. & Webster, C. Cytoplasmic activation of human nuclear genes in stable heterocaryons. *Cell* **32**, 1171-1180 (1983).
- 43 Chiu, C. P. & Blau, H. M. Reprogramming cell differentiation in the absence of DNA synthesis. *Cell* **37**, 879-887 (1984).
- 44 Blau, H. M. *et al.* Plasticity of the differentiated state. *Science* **230**, 758-766 (1985).
- 45 Ieda, M. *et al.* Direct reprogramming of fibroblasts into functional cardiomyocytes by defined factors. *Cell* **142**, 375-386, doi:10.1016/j.cell.2010.07.002 (2010).
- 46 Margariti, A. *et al.* Direct reprogramming of fibroblasts into endothelial cells capable of angiogenesis and reendothelialization in tissue-engineered vessels. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 13793-13798, doi:10.1073/pnas.1205526109 (2012).
- 47 Zhu, J. *et al.* Genome-wide chromatin state transitions associated with developmental and environmental cues. *Cell* **152**, 642-654, doi:10.1016/j.cell.2012.12.033 (2013).
- 48 Yadav, T., Quivy, J. P. & Almouzni, G. Chromatin plasticity: A versatile landscape that underlies cell fate and identity. *Science* **361**, 1332-+, doi:10.1126/science.aat8950 (2018).

- 49 Barrero, M. J., Boue, S. & Belmonte, J. C. I. Epigenetic Mechanisms that Regulate Cell Identity. *Cell stem cell* **7**, 565-570, doi:10.1016/j.stem.2010.10.009 (2010).
- 50 Theunissen, T. W. & Jaenisch, R. Molecular control of induced pluripotency. *Cell stem cell* **14**, 720-734, doi:10.1016/j.stem.2014.05.002 (2014).
- 51 Apostolou, E. & Hochedlinger, K. Chromatin dynamics during cellular reprogramming. *Nature* **502**, 462-471, doi:10.1038/nature12749 (2013).
- 52 Papp, B. & Plath, K. Epigenetics of reprogramming to induced pluripotency. *Cell* **152**, 1324-1343, doi:10.1016/j.cell.2013.02.043 (2013).
- 53 Seisenberger, S. *et al.* Reprogramming DNA methylation in the mammalian life cycle: building and breaking epigenetic barriers. *Philos Trans R Soc Lond B Biol Sci* **368**, 20110330, doi:10.1098/rstb.2011.0330 (2013).
- 54 Messerschmidt, D. M., Knowles, B. B. & Solter, D. DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. *Genes & development* **28**, 812-828, doi:10.1101/gad.234294.113 (2014).
- 55 Seisenberger, S., Peat, J. R. & Reik, W. Conceptual links between DNA methylation reprogramming in the early embryo and primordial germ cells. *Curr Opin Cell Biol* **25**, 281-288, doi:10.1016/j.ceb.2013.02.013 (2013).
- 56 Broske, A. M. *et al.* DNA methylation protects hematopoietic stem cell multipotency from myeloid restriction. *Nat Genet* **41**, 1207-1215, doi:10.1038/ng.463 (2009).
- 57 Chen, G. *et al.* Single-cell analyses of X Chromosome inactivation dynamics and pluripotency during differentiation. *Genome Res* **26**, 1342-1354, doi:10.1101/gr.201954.115 (2016).
- 58 De Carvalho, D. D., You, J. S. & Jones, P. A. DNA methylation and cellular reprogramming. *Trends Cell Biol* **20**, 609-617, doi:10.1016/j.tcb.2010.08.003 (2010).
- 59 Looney, T. J. *et al.* Systematic mapping of occluded genes by cell fusion reveals prevalence and stability of cis-mediated silencing in somatic cells. *Genome Res* **24**, 267-280, doi:10.1101/gr.143891.112 (2014).
- 60 Chiu, C. P. & Blau, H. M. 5-Azacytidine permits gene activation in a previously noninducible cell type. *Cell* **40**, 417-424 (1985).
- 61 Mikkelsen, T. S. *et al.* Dissecting direct reprogramming through integrative genomic analysis. *Nature* **454**, 49-55, doi:10.1038/nature07056 (2008).
- 62 Bhutani, N. *et al.* Reprogramming towards pluripotency requires AID-dependent DNA demethylation. *Nature* **463**, 1042-1047, doi:10.1038/nature08752 (2010).
- 63 Rais, Y. *et al.* Deterministic direct reprogramming of somatic cells to pluripotency. *Nature* **502**, 65-70, doi:10.1038/nature12587 (2013).
- 64 Lee, D. S. *et al.* An epigenomic roadmap to induced pluripotency reveals DNA methylation as a reprogramming modulator. *Nature communications* **5**, 5619, doi:10.1038/ncomms6619 (2014).

- 65 Buschbeck, M. & Hake, S. B. Variants of core histones and their roles in cell fate decisions, development and cancer. *Nature reviews. Molecular cell biology* **18**, 299-314, doi:10.1038/nrm.2016.166 (2017).
- 66 Pasque, V. *et al.* Histone variant macroH2A marks embryonic differentiation in vivo and acts as an epigenetic barrier to induced pluripotency. *Journal of cell science* **125**, 6094-6104, doi:10.1242/jcs.113019 (2012).
- 67 Almouzni, G. & Wolffe, A. P. Replication-coupled chromatin assembly is required for the repression of basal transcription in vivo. *Genes & development* **7**, 15 (1993).
- 68 Filion, G. J. *et al.* Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. *Cell* **143**, 212-224, doi:10.1016/j.cell.2010.09.009 (2010).
- 69 Benayoun, B. A. *et al.* H3K4me3 breadth is linked to cell identity and transcriptional consistency. *Cell* **158**, 673-688, doi:10.1016/j.cell.2014.06.027 (2014).
- 70 Becker, J. S., Nicetto, D. & Zaret, K. S. H3K9me3-Dependent Heterochromatin: Barrier to Cell Fate Changes. *Trends Genet* **32**, 29-41, doi:10.1016/j.tig.2015.11.001 (2016).
- 71 Matoba, S. *et al.* Embryonic development following somatic cell nuclear transfer impeded by persisting histone methylation. *Cell* **159**, 884-895, doi:10.1016/j.cell.2014.09.055 (2014).
- 72 Becker, J. S. *et al.* Genomic and Proteomic Resolution of Heterochromatin and Its Restriction of Alternate Fate Genes. *Molecular cell* **68**, 1023-1037 e1015, doi:10.1016/j.molcel.2017.11.030 (2017).
- 73 Chen, T. & Dent, S. Y. Chromatin modifiers and remodellers: regulators of cellular differentiation. *Nature reviews. Genetics* **15**, 93-106, doi:10.1038/nrg3607 (2014).
- 74 Onder, T. T. *et al.* Chromatin-modifying enzymes as modulators of reprogramming. *Nature* **483**, 598-U119, doi:10.1038/nature10953 (2012).
- 75 Cheloufi, S. *et al.* The histone chaperone CAF-1 safeguards somatic cell identity. *Nature* **528**, 218-224, doi:10.1038/nature15749 (2015).
- 76 Ishiuchi, T. *et al.* Early embryonic-like cells are induced by downregulating replication-dependent chromatin assembly. *Nat Struct Mol Biol* **22**, 662-671, doi:10.1038/nsmb.3066 (2015).
- 77 Qin, H. *et al.* Systematic identification of barriers to human iPSC generation. *Cell* **158**, 449-461, doi:10.1016/j.cell.2014.05.040 (2014).
- 78 Winzi, M., Paszkowski-Rogacz, M. & Buchholz, F. Another Brick in the Wall: RNAi Screens Identify New Barriers in iPSC Reprogramming. *Cell stem cell* **15**, 116-118, doi:10.1016/j.stem.2014.07.008 (2014).
- 79 Yang, C. S., Chang, K. Y. & Rana, T. M. Genome-wide functional analysis reveals factors needed at the transition steps of induced reprogramming. *Cell Rep* **8**, 327-337, doi:10.1016/j.celrep.2014.07.002 (2014).
- 80 Egli, D., Birkhoff, G. & Eggan, K. Mediators of reprogramming: transcription factors and transitions through mitosis. *Nature reviews. Molecular cell biology* **9**, 505-516, doi:10.1038/nrm2439 (2008).
- 81 Xu, Y. *et al.* Transcriptional Control of Somatic Cell Reprogramming. *Trends Cell Biol* **26**, 272-288, doi:10.1016/j.tcb.2015.12.003 (2016).

- 82 Thomas, S. *et al.* Dynamic reprogramming of chromatin accessibility during *Drosophila* embryo development. *Genome biology* **12**, R43, doi:10.1186/gb-2011-12-5-r43 (2011).
- 83 Kaplan, T. *et al.* Quantitative models of the mechanisms that control genome-wide patterns of transcription factor binding during early *Drosophila* development. *PLoS genetics* **7**, e1001290, doi:10.1371/journal.pgen.1001290 (2011).
- 84 Li, X. Y. *et al.* The role of chromatin accessibility in directing the widespread, overlapping patterns of *Drosophila* transcription factor binding. *Genome biology* **12**, R34, doi:10.1186/gb-2011-12-4-r34 (2011).
- 85 Adamson, E. D. & Woodland, H. R. Histone synthesis in early amphibian development: histone and DNA syntheses are not co-ordinated. *Journal of molecular biology* **88**, 263-285 (1974).
- 86 Amodeo, A. A., Jukam, D., Straight, A. F. & Skotheim, J. M. Histone titration against the genome sets the DNA-to-cytoplasm threshold for the *Xenopus* midblastula transition. *Proceedings of the National Academy of Sciences of the United States of America* **112**, E1086-1095, doi:10.1073/pnas.1413990112 (2015).
- 87 Joseph, S. R. *et al.* Competition between histone and transcription factor binding regulates the onset of transcription in zebrafish embryos. *Elife* **6**, doi:10.7554/eLife.23326 (2017).
- 88 Niwa, H., Miyazaki, J. & Smith, A. G. Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet* **24**, 372-376, doi:10.1038/74199 (2000).
- 89 Kopp, J. L., Ormsbee, B. D., Desler, M. & Rizzino, A. Small increases in the level of Sox2 trigger the differentiation of mouse embryonic stem cells. *Stem cells* **26**, 903-911, doi:10.1634/stemcells.2007-0951 (2008).
- 90 Nishimura, K. *et al.* Manipulation of KLF4 expression generates iPSCs paused at successive stages of reprogramming. *Stem Cell Reports* **3**, 915-929, doi:10.1016/j.stemcr.2014.08.014 (2014).
- 91 Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**, 1917-1920, doi:10.1126/science.1151526 (2007).
- 92 Vivien, C. *et al.* Non-viral expression of mouse Oct4, Sox2, and Klf4 transcription factors efficiently reprograms tadpole muscle fibers in vivo. *The Journal of biological chemistry* **287**, 7427-7435, doi:10.1074/jbc.M111.324368 (2012).
- 93 Hanna, J. *et al.* Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature* **462**, 595-601, doi:10.1038/nature08592 (2009).
- 94 Schmidt, R. & Plath, K. The roles of the reprogramming factors Oct4, Sox2 and Klf4 in resetting the somatic cell epigenome during induced pluripotent stem cell generation. *Genome biology* **13**, 251, doi:10.1186/gb-2012-13-10-251 (2012).
- 95 Chronis, C. *et al.* Cooperative Binding of Transcription Factors Orchestrates Reprogramming. *Cell* **168**, 442-459 e420, doi:10.1016/j.cell.2016.12.016 (2017).

- 96 Smale, S. T. Pioneer factors in embryonic stem cells and differentiation. *Current opinion in genetics & development* **20**, 519-526, doi:10.1016/j.gde.2010.06.010 (2010).
- 97 Zaret, K. S. Regulatory phases of early liver development: paradigms of organogenesis. *Nature reviews. Genetics* **3**, 499-512, doi:10.1038/nrg837 (2002).
- 98 Lee, C. S., Friedman, J. R., Fulmer, J. T. & Kaestner, K. H. The initiation of liver development is dependent on Foxa transcription factors. *Nature* **435**, 944-947, doi:10.1038/nature03649 (2005).
- 99 Sekiya, T., Muthurajan, U. M., Luger, K., Tulin, A. V. & Zaret, K. S. Nucleosome-binding affinity as a primary determinant of the nuclear mobility of the pioneer transcription factor FoxA. *Genes & development* **23**, 804-809, doi:10.1101/gad.1775509 (2009).
- 100 Swinstead, E. E. *et al.* Steroid Receptors Reprogram FoxA1 Occupancy through Dynamic Chromatin Transitions. *Cell* **165**, 593-605, doi:10.1016/j.cell.2016.02.067 (2016).
- 101 Vierbuchen, T. *et al.* Direct conversion of fibroblasts to functional neurons by defined factors. *Nature* **463**, 1035-1041, doi:10.1038/nature08797 (2010).
- 102 Marro, S. *et al.* Direct lineage conversion of terminally differentiated hepatocytes to functional neurons. *Cell stem cell* **9**, 374-382, doi:10.1016/j.stem.2011.09.002 (2011).
- 103 Wapinski, O. L. *et al.* Hierarchical mechanisms for direct reprogramming of fibroblasts to neurons. *Cell* **155**, 621-635, doi:10.1016/j.cell.2013.09.028 (2013).
- 104 Soufi, A., Donahue, G. & Zaret, K. S. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell* **151**, 994-1004, doi:10.1016/j.cell.2012.09.045 (2012).
- 105 Soufi, A. *et al.* Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell* **161**, 555-568, doi:10.1016/j.cell.2015.03.017 (2015).
- 106 Lowry, W. E. Does transcription factor induced pluripotency accurately mimic embryo derived pluripotency? *Current opinion in genetics & development* **22**, 429-434, doi:10.1016/j.gde.2012.07.003 (2012).
- 107 Li, M. & Belmonte, J. C. Ground rules of the pluripotency gene regulatory network. *Nature reviews. Genetics* **18**, 180-191, doi:10.1038/nrg.2016.156 (2017).
- 108 Niwa, H. How is pluripotency determined and maintained? *Development* **134**, 635-646, doi:10.1242/dev.02787 (2007).
- 109 Do, J. T. *et al.* Erasure of cellular memory by fusion with pluripotent cells. *Stem cells* **25**, 1013-1020, doi:10.1634/stemcells.2006-0691 (2007).
- 110 Nakatake, Y. *et al.* Klf4 cooperates with Oct3/4 and Sox2 to activate the Lefty1 core promoter in embryonic stem cells. *Molecular and cellular biology* **26**, 7772-7782, doi:10.1128/MCB.00468-06 (2006).
- 111 Boer, B. *et al.* Elevating the levels of Sox2 in embryonal carcinoma cells and embryonic stem cells inhibits the expression of Sox2:Oct-3/4 target genes. *Nucleic acids research* **35**, 1773-1786, doi:10.1093/nar/gkm059 (2007).

- 112 Ng, H. H. & Surani, M. A. The transcriptional and signalling networks of pluripotency. *Nature cell biology* **13**, 490-496, doi:10.1038/ncb0511-490 (2011).
- 113 Hall, J. *et al.* Oct4 and LIF/Stat3 additively induce Kruppel factors to sustain embryonic stem cell self-renewal. *Cell stem cell* **5**, 597-609, doi:10.1016/j.stem.2009.11.003 (2009).
- 114 Yeo, J. C. *et al.* Klf2 is an essential factor that sustains ground state pluripotency. *Cell stem cell* **14**, 864-872, doi:10.1016/j.stem.2014.04.015 (2014).
- 115 Tetreault, M. P., Yang, Y. & Katz, J. P. Kruppel-like factors in cancer. *Nature reviews. Cancer* **13**, 701-713, doi:10.1038/nrc3582 (2013).
- 116 Frankenberg, S. R. *et al.* The POU-er of gene nomenclature. *Development* **141**, 2921-2923, doi:10.1242/dev.108407 (2014).
- 117 Abdelalim, E. M., Emara, M. M. & Kolatkar, P. R. The SOX Transcription Factors as Key Players in Pluripotent Stem Cells. *Stem Cells and Development* **23**, 2687-2699, doi:10.1089/scd.2014.0297 (2014).
- 118 Jiang, J. *et al.* A core Klf circuitry regulates self-renewal of embryonic stem cells. *Nature cell biology* **10**, 353-360, doi:10.1038/ncb1698 (2008).
- 119 Bourillot, P. Y. & Savatier, P. Kruppel-like transcription factors and control of pluripotency. *BMC Biol* **8**, 125, doi:10.1186/1741-7007-8-125 (2010).
- 120 Parisi, S. *et al.* Direct targets of Klf5 transcription factor contribute to the maintenance of mouse embryonic stem cell undifferentiated state. *BMC Biol* **8**, 128, doi:10.1186/1741-7007-8-128 (2010).
- 121 Long, X. & Singla, D. K. Inactivation of Klf5 by zinc finger nuclease downregulates expression of pluripotent genes and attenuates colony formation in embryonic stem cells. *Mol Cell Biochem* **382**, 113-119, doi:10.1007/s11010-013-1724-5 (2013).
- 122 Aksoy, I. *et al.* Klf4 and Klf5 differentially inhibit mesoderm and endoderm differentiation in embryonic stem cells. *Nature communications* **5**, 3719, doi:10.1038/ncomms4719 (2014).
- 123 Aksoy, I. *et al.* Sox transcription factors require selective interactions with Oct4 and specific transactivation functions to mediate reprogramming. *Stem cells* **31**, 2632-2646, doi:10.1002/stem.1522 (2013).
- 124 Hinkley, C. S., Martin, J. F., Leibham, D. & Perry, M. Sequential expression of multiple POU proteins during amphibian early development. *Molecular and cellular biology* **12**, 638-649 (1992).
- 125 Morichika, K., Sugimoto, M., Yasuda, K. & Kinoshita, T. Possible regulation of Oct60 transcription by a positive feedback loop in *Xenopus* oocytes. *Zygote* **22**, 266-274, doi:10.1017/S0967199412000536 (2014).
- 126 Young, J. J., Kjolby, R. A., Kong, N. R., Monica, S. D. & Harland, R. M. Spalt-like 4 promotes posterior neural fates via repression of pou5f3 family members in *Xenopus*. *Development* **141**, 1683-1693, doi:10.1242/dev.099374 (2014).
- 127 Liu, S. B., G; Sun, R.;Guo, S.; Xue, B.; Wei, R.; Cooney, A. J.; Liu, Z. Sox2 is the faithful marker for pluripotency in pig: evidence from

- embryonic studies. *Developmental dynamics : an official publication of the American Association of Anatomists* **244**, 9 (2015).
- 128 Cao, Q. *et al.* Klf4 is required for germ-layer differentiation and body axis patterning during *Xenopus* embryogenesis. *Development* **139**, 3950-3961, doi:10.1242/dev.082024 (2012).
- 129 Gao, Y. *et al.* Kruppel-like factor family genes are expressed during *Xenopus* embryogenesis and involved in germ layer formation and body axis patterning. *Developmental dynamics : an official publication of the American Association of Anatomists* **244**, 1328-1346, doi:10.1002/dvdy.24310 (2015).
- 130 Schmidt, D. *et al.* Five-vertebrate ChIP-seq reveals the evolutionary dynamics of transcription factor binding. *Science* **328**, 1036-1040, doi:10.1126/science.1186176 (2010).
- 131 Tapia, N. *et al.* Reprogramming to pluripotency is an ancient trait of vertebrate Oct4 and Pou2 proteins. *Nature communications* **3**, doi:ARTN 1279 10.1038/ncomms2229 (2012).
- 132 Morrison, G. M. & Brickman, J. M. Conserved roles for Oct4 homologues in maintaining multipotency during early vertebrate development. *Development* **133**, 2011-2022, doi:10.1242/dev.02362 (2006).
- 133 Hammachi, F. *et al.* Transcriptional activation by Oct4 is sufficient for the maintenance and induction of pluripotency. *Cell Rep* **1**, 99-109, doi:10.1016/j.celrep.2011.12.002 (2012).
- 134 Dixon, J. E. *et al.* Axolotl Nanog activity in mouse embryonic stem cells demonstrates that ground state pluripotency is conserved from urodele amphibians to mammals. *Development* **137**, 2973-2980, doi:10.1242/dev.049262 (2010).
- 135 Halley-Stott, R. P. *et al.* Mammalian nuclear transplantation to Germinal Vesicle stage *Xenopus* oocytes - a method for quantitative transcriptional reprogramming. *Methods* **51**, 56-65, doi:10.1016/j.ymeth.2010.01.035 (2010).
- 136 Jullien, J. *et al.* Hierarchical molecular events driven by oocyte-specific factors lead to rapid and extensive reprogramming. *Molecular cell* **55**, 524-536, doi:10.1016/j.molcel.2014.06.024 (2014).
- 137 Roberts, T. C. *et al.* Quantification of nascent transcription by bromouridine immunocapture nuclear run-on RT-qPCR. *Nature protocols* **10**, 1198-1211, doi:10.1038/nprot.2015.076 (2015).
- 138 Oka, M., Moriyama, T., Asally, M., Kawakami, K. & Yoneda, Y. Differential Role for Transcription Factor Oct4 Nucleocytoplasmic Dynamics in Somatic Cell Reprogramming and Self-renewal of Embryonic Stem Cells. *Journal of Biological Chemistry* **288**, 15085-15097, doi:10.1074/jbc.M112.448837 (2013).
- 139 Sutinen, P., Rahkama, V., Rytinki, M. & Palvimo, J. J. Nuclear mobility and activity of FOXA1 with androgen receptor are regulated by SUMOylation. *Mol Endocrinol* **28**, 1719-1728, doi:10.1210/me.2014-1035 (2014).
- 140 Yanai, I., Peshkin, L., Jorgensen, P. & Kirschner, M. W. Mapping gene expression in two *Xenopus* species: evolutionary constraints and developmental flexibility. *Developmental cell* **20**, 483-496, doi:10.1016/j.devcel.2011.03.015 (2011).

- 141 Peshkin, L. *et al.* On the Relationship of Protein and mRNA Dynamics in Vertebrate Embryonic Development. *Developmental cell* **35**, 383-394, doi:10.1016/j.devcel.2015.10.010 (2015).
- 142 Sun, L. *et al.* Quantitative proteomics of *Xenopus laevis* embryos: expression kinetics of nearly 4000 proteins during early development. *Sci Rep* **4**, 4365, doi:10.1038/srep04365 (2014).
- 143 Wingender, E. The TRANSFAC project as an example of framework technology that supports the analysis of genomic regulation. *Brief Bioinform* **9**, 326-332, doi:10.1093/bib/bbn016 (2008).
- 144 Eisenberg, E. & Levanon, E. Y. Human housekeeping genes, revisited. *Trends Genet* **29**, 569-574, doi:10.1016/j.tig.2013.05.010 (2013).
- 145 Biddle, A., Simeoni, I. & Gurdon, J. B. *Xenopus* oocytes reactivate muscle gene transcription in transplanted somatic nuclei independently of myogenic factors. *Development* **136**, 2695-2703, doi:10.1242/dev.036327 (2009).
- 146 Owens, N. D. *et al.* Measuring Absolute RNA Copy Numbers at High Temporal Resolution Reveals Transcriptome Kinetics in Development. *Cell Rep* **14**, 632-647, doi:10.1016/j.celrep.2015.12.050 (2016).
- 147 Smits, A. H. *et al.* Global absolute quantification reveals tight regulation of protein expression in single *Xenopus* eggs. *Nucleic acids research* **42**, 9880-9891, doi:10.1093/nar/gku661 (2014).
- 148 Wuhr, M. *et al.* Deep proteomics of the *Xenopus laevis* egg using an mRNA-derived reference database. *Current biology : CB* **24**, 1467-1475, doi:10.1016/j.cub.2014.05.044 (2014).
- 149 Iwafuchi-Doi, M. & Zaret, K. S. Cell fate control by pioneer transcription factors. *Development* **143**, 1833-1837, doi:10.1242/dev.133900 (2016).
- 150 Metzakopian, E. *et al.* Genome-wide characterisation of Foxa1 binding sites reveals several mechanisms for regulating neuronal differentiation in midbrain dopamine cells. *Development* **142**, 1315-1324, doi:10.1242/dev.115808 (2015).
- 151 Metzakopian, E. *et al.* Genome-wide characterization of Foxa2 targets reveals upregulation of floor plate genes and repression of ventrolateral genes in midbrain dopaminergic progenitors. *Development* **139**, 2625-2634, doi:10.1242/dev.081034 (2012).
- 152 Belaguli, N. S., Zhang, M., Brunicardi, F. C. & Berger, D. H. Forkhead box protein A2 (FOXA2) protein stability and activity are regulated by sumoylation. *PloS one* **7**, e48019, doi:10.1371/journal.pone.0048019 (2012).
- 153 Gurdon, J. B. Changes in somatic cell nuclei inserted into growing and maturing amphibian oocytes. *J Embryol Exp Morphol* **20**, 401-414 (1968).
- 154 Jullien, J., Astrand, C., Halley-Stott, R. P., Garrett, N. & Gurdon, J. B. Characterization of somatic cell nuclear reprogramming by oocytes in which a linker histone is required for pluripotency gene reactivation. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 5483-5488, doi:10.1073/pnas.1000599107 (2010).
- 155 Narbonne, P., Simpson, D. E. & Gurdon, J. B. Deficient induction response in a *Xenopus* nucleocytoplasmic hybrid. *PLoS biology* **9**, e1001197, doi:10.1371/journal.pbio.1001197 (2011).

- 156 Okada, Y. Analysis of giant polynuclear cell formation caused by HVJ virus from Ehrlich's ascites tumor cells. I. Microscopic observation of giant polynuclear cell formation. *Experimental cell research* **26**, 98-107 (1962).
- 157 Okada, Y. Analysis of giant polynuclear cell formation caused by HVJ virus from Ehrlich's ascites tumor cells. III. Relationship between cell condition and fusion reaction or cell degeneration reaction. *Experimental cell research* **26**, 119-128 (1962).
- 158 Okada, Y. & Tadokoro, J. Analysis of giant polynuclear cell formation caused by HVJ virus from Ehrlich's ascites tumor cells. II. Quantitative analysis of giant polynuclear cell formation. *Experimental cell research* **26**, 108-118 (1962).
- 159 Harris, H. The reactivation of the red cell nucleus. *Journal of cell science* **2**, 23-32 (1967).
- 160 Harris, H., Miller, O. J., Klein, G., Worst, P. & Tachibana, T. Suppression of malignancy by cell fusion. *Nature* **223**, 363-368 (1969).
- 161 Tsubouchi, T. *et al.* DNA synthesis is required for reprogramming mediated by stem cell fusion. *Cell* **152**, 873-883, doi:10.1016/j.cell.2013.01.012 (2013).
- 162 Pereira, C. F. *et al.* Heterokaryon-based reprogramming of human B lymphocytes for pluripotency requires Oct4 but not Sox2. *PLoS genetics* **4**, e1000170, doi:10.1371/journal.pgen.1000170 (2008).
- 163 Terranova, R., Pereira, C. F., Du Roure, C., Merckenschlager, M. & Fisher, A. G. Acquisition and extinction of gene expression programs are separable events in heterokaryon reprogramming. *Journal of cell science* **119**, 2065-2072, doi:10.1242/jcs.02945 (2006).
- 164 Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 3 (2011).
- 165 Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105-1111, doi:10.1093/bioinformatics/btp120 (2009).
- 166 Anders, S., Pyl, P. T. & Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166-169, doi:10.1093/bioinformatics/btu638 (2015).
- 167 McCarthy, D. J., Chen, Y. & Smyth, G. K. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic acids research* **40**, 4288-4297, doi:10.1093/nar/gks042 (2012).

Chapter 10 Appendices

10.1 Appendix I: Plasmid construction

To overexpress TF proteins in *Xenopus* oocytes, the TF mRNA was made via *in vitro* transcription. First, the coding sequence of TF on plasmids was amplified by PCR and inserted into pTOPO vector via Gateway cloning system. However, while I constructed the plasmids, there was a missense mutation found on the coding sequence of mFoxa1. Therefore, the mutagenesis PCR was performed to fix the mutation and the fixed mFoxa1 was inserted into pCS2-HA vector to have T7 promoter on plasmid and add HA tags to mFoxa1. The TF mRNA was made from the TF containing plasmids via *in vitro* transcription.

10.1.1 Construction of pTOPO-mFoxa1 plasmid from pBabe-puro-FoxA1 plasmid and pTOPO vector via Gateway cloning system

The coding sequence of mFoxa1 on pBabe-puro-FoxA1 plasmid was first amplified by high-fidelity DNA polymerase (Phusion High-Fidelity DNA polymerase, 2U/μl, Cat No M0530, New England Biolabs) with primers designed for Gateway cloning system (Figure 10.1.1). The amount of reagents and cycling parameter for PCR is as follows:

1μl 100ng/μl plasmid (Miniprep eluent), 32.5μl H₂O, 10μl 5X Phusion HF Buffer (NEB), 1μl dNTPs (10mM each, NEB), 2.5μl Forward primer (10μM, Sigma-Aldrich), 2.5μl Reverse primer (10μM, Sigma-Aldrich), 0.5μl Phusion DNA polymerase (NEB)

Number of cycles	Temperature	Duration
1	98°C	30 secs
26	98°C	10 secs
	62°C	30 secs
	72°C	2 mins
1	72°C	10 mins
1	4	∞

The mFoxa1 amplicons were then inserted into pENTR/D-TOPO vector and introduced into competent *E. coli* via heat-shock transformation (pENTR/D-TOPO cloning kit with One Shot TOP10 Chemically Competent *E. coli*, Cat No K240020, Thermo Fisher Scientific). The transformed *E. coli* was then selected by kanamycin and the propagated plasmids were extracted (QIAprep Spin Miniprep Kit, QIAGEN) and sequenced (Department of Biochemistry, University of Cambridge). However, the sequencing results showed there was a missense mutation on the coding sequence of mFoxa1 and it was also found on the original pBabe-puro-FoxA1 so the mutagenesis was utilized to fix it.

Figure 10.1.1



Figure 10.1.1 Primers for amplifying coding sequence of mFoxa1 on pBabe-puro-FoxA1

10.1.2 Fixation of the missense mutation of mFoxa1 coding sequence on pTOPO-mFoxa1 plasmid

Since the mFoxa1 coding sequence of pBabe-puro-FoxA1 plasmid was spotted with a missense mutation from Ser-12 to Asn-12, the mutation of plasmid was firstly mutated back to the original Ser-12 by using designed primers (Figure 10.1.2). New plasmids containing the designed primers was synthesized by using high-fidelity DNA polymerase (*PfuTurbo* DNA polymerase, 2.5U/ μ l, Cat No 600252, Agilent Technologies) with the PCR cycling parameters:

2 μ l pTOPO-mFoxa1 plasmid (Miniprep eluent), 5 μ l 10X Cloned *Pfu* buffer (Agilent), 1.25 μ l Forward primer (100 μ M, Sigma-Aldrich), 1.25 μ g 100 μ g/ μ l Reverse primer (100 μ M, Sigma-Aldrich), 1 μ l dNTPs (10mM each), 38.5 μ l H₂O, 1 μ l *PfuTurbo* DNA polymerase (Agilent)

Number of cycles	Temperature	Duration
1	95°C	30 secs
12	95°C	30 secs
	55°C	1 min
	68°C	8 mins
1	4°C	∞

The amplicons were introduced into competent *E. coli* via heat-shock transformation (One Shot TOP10 Chemically Competent *E. coli*, Thermo Fisher Scientific). The propagated plasmids were then extracted (QIAprep Spin Miniprep Kit, QIAGEN) and the fixed sequence was confirmed by Sanger sequencing (Department of Biochemistry, University of Cambridge).

Figure 10.1.2

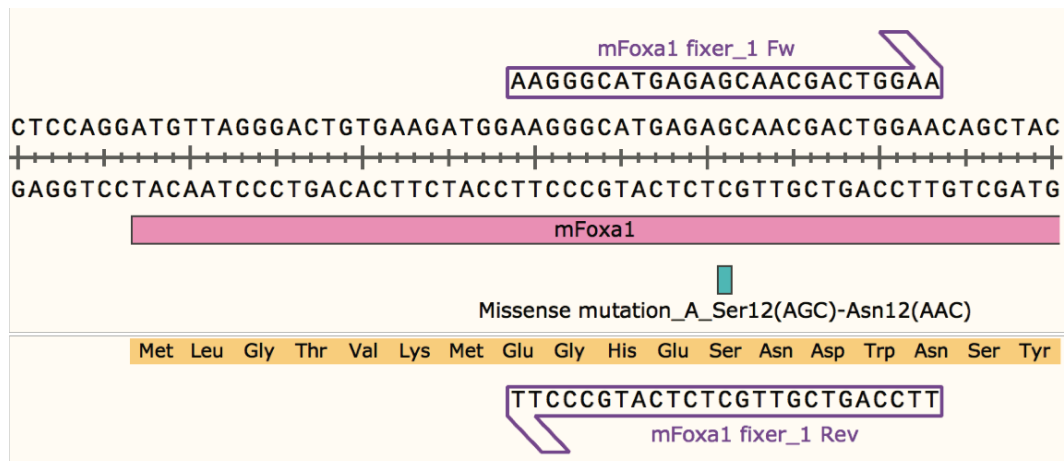


Figure 10.1.2 Missense mutation of pBabe-puro-FoxA1 plasmid and primers used for fixing the plasmid

10.1.3 Construction of pCS2-mFoxa1-3HA plasmid from pTOPO-mFoxa1 plasmid and pCS2-3HA-attR vector

To add HA tags to the N-terminus of mFoxa1, the mFoxa1 coding sequence on pTOPO-mFoxa1 plasmid was recombined with pCS2-3HA-attR vector via LR clonase (Gateway LR clonase II enzyme mix, Thermo Fisher Scientific).

The reaction products were then introduced into competent *E. coli* via heat-shock transformation (One Shot TOP10 Chemically Competent *E. coli*, Thermo Fisher Scientific) and the successful transformed *E. coli* was selected by Ampicillin. The resulting propagated plasmids were then extracted (QIAprep Spin Miniprep Kit, QIAGEN) and sequenced (Department of Biochemistry, University of Cambridge).

10.2 Appendix II: Primer list for neurogenic genes

Gene Symbol	Forward primer	Reverse Primer
Lmx1b	GTGGAGATGACGGGAAAGAC	CTTGAAAGCTCTTCGCTGCT
Dmrtb1	GGCTACCTCTCTGCTCTCCA	GCTTCTGCATCCTGGTCATT
Lmcd1	TACCAGCGTGTGGAAGACCT	TGACCCTTGGTGACGATGTA
Smarca1	GGCTCAAATTGAACGTGGAG	ATCTTGCAATTTTGGCATCC
Arx	TTTTCTAGGAGCAGCGGTGT	GGGGCCATAGTGGAAGAG
Kcnip3	GGCCATCATGAAGTCCATCT	AACCTCTCCACATGCTCCAG
Elk3	TGGAGCAGCCTTAGTCCTGT	GTTGAGCAGTGTGGGGAAGT
Rora	CGTGGCTTCAGGAAAAGGTA	TTCCATCTTCTCGGTGGTTC
Ferd3l	CAGAAAGGTACCCACCTTCG	GAGCTCGGTCATGAAGGAAA
Gli1	TCAGCTGGACTTTGTGGCTA	AGAGGGAGATGGGGTGTTTT
Otx2	AGTCCCCAGCTTCTCTTTCC	TTCCAAGAGGCAGTTTGGTC
Foxb1	CCACCCTGCTCTCGAACTC	CTGGTGAGTGTCTCGCTAGG
Meis2	TGACCAGTCAAATCGAGCAG	ATGTGTTGCTGACCATCCAA
Otx1	CATCATCACCATCACCTCA	CAGAGCCTCCATAACCTTGG

10.3 Appendix III: Bioinformatics analysis

10.3.1 Hybrid genome generation

Fasta files from *Mus musculus* (mm10) and *Xenopus laevis* (xla9.1) have been concatenated one after the other in order to create a hybrid large *mouse-Xenopus* genome. To distinguish the *X. laevis* chromosomes in the fasta file, they have been renamed as "xla_chr" instead of just "chr". Similarly, the gtf files containing the annotation of all transcripts from Mouse (mm10) and of all primary transcript from *X. Laevis* (xla 9.1) have been concatenated.

10.3.2 Processing of fastq files

FastQ files were processed with *cutadapt* (version 1.9.1, options *-q 10 -O 3*)¹⁶⁴ for adapter trimming. Filtered reads were then aligned to the hybrid *mouse-Xenopus* genome with *tophat* (version v2.1.1)¹⁶⁵. Transcripts were assigned to gene and counted using *htseq-count* (HTSeq-0.5.4p3)¹⁶⁶.

10.3.3 Reproducibility analysis

We performed hierarchical clustering of the libraries by using the *hclust* function of the *stats* package in R with default parameters. Clustering has been performed on both raw counts and then on normalized data. Data have been normalized with a double scaling procedure that uses the z-score transformation. The first scaling has been done on experiments belonging to the same batch (i.e. experiments produced at the same time) in order to reduce the batch variability. Then they have been scaled again.

Multidimensional scaling plots were obtained through the *plotMDS* function of the *limma* package in *bioconductor* using the normalized data.

10.3.4 Differential expression (DE) analysis.

Data were normalized and analyzed for differential expression using the package *edgeR*¹⁶⁷ from *Bioconductor*. The functions *glmFit* and *glmLRT* were used to assess differential expression (logarithm on base 2 of fold change, logFC) and significance (false discovery rate, FDR) between each pair of conditions we explored. The DE analysis was performed considering how the experiments were paired by creating a design matrix with *model.matrix* from the *stats* package in order to keep into account the information of the frog, the treatment and the replicate. The set of genes used for the DE analysis was obtained by filtering those with at least 1 cpm (count per million) in half (50%) of the total number of the experiments under investigation. Genes differentially expressed significantly between samples was determined considering those with FDR less than 0.1. No cutoff is applied for the threshold of the value of log₂FC.

10.4 Appendix IV: Preliminary data validation for reprogrammed transcriptomes of mESCs, MEFs and mMyos by *Xenopus* oocytes

As negative controls, Oocyte-NT samples without BrUTP injection were prepared in parallel with Oocyte-NT samples with BrUTP injection (n=2, Frog 2 and 6, Table 10.4.1). In MEF libraries of Frog 2, the mapped reads/nM increased from 2 million to 13 million reads/nM for *Mus musculus* and decreased from 27 million to 8 million reads/nM for *Xenopus laevis* after BrUTP pull-down (Frog 2, #2, Table 10.4.1). Likewise, the mapped reads/nM increase from 1 to 6 million reads/nM for *Mus musculus* and decrease from 11 to 4 million reads/nM for *Xenopus laevis* in Frog 6 after BrUTP pull-down (Frog 6, #6 and #7, Table 10.4.1). Therefore, the average ratio of mapped reads of *Mus musculus* to *Xenopus laevis* is 1:13 for no BrUTP samples (2:27 for Frog 2 and 1:11 for Frog 6) and 1.6:1 for BrUTP samples (13:8 for Frog 2 and 6:7 for Frog 6) after BrUTP pull-down.

This indicates that genes of *Mus musculus* are actively transcribed after Oocyte-NT and BrUTP pull-down can efficiently isolate newly synthesized transcripts, mostly from *Mus musculus*. Furthermore, most non-specifically bound transcripts are from *Xenopus laevis*, as judged by the ratio of mapped reads in no BrUTP samples, and those non-specifically bound transcripts can be disregarded effectively after BrUTP pull-down.

Interestingly, the mapped reads/nM for *Mus musculus* of MEF-NT plus BrUTP samples in Frog 1 and Frog 2 do not increase from Day 1 to Day 2 after

Oocyte-NT (Table 10.4.1). Additionally, the average ratio of mapped reads/nM of mESC-NT plus BrUTP to mMyo-NT plus BrUTP in Frog 2, 3 and 4 is 4 to 1 (Table 10.4.1). It suggests that the amount of newly synthesized transcripts in MEF-NT keeps steady 24 hours after Oocyte-NT for the balance between synthesis and degradation of polyadenylated RNA. Furthermore, there are more newly synthesized transcripts from nuclei of mESCs than mMyos possibly because the chromatin of mESCs is generally more accessible to the effect of oocyte factors than the genes of mMyos⁴⁸.

For an overview of all the RNA-seq libraries used in this chapter, hierarchical clustering was applied and libraries are clustered by the expression level of newly synthesized transcripts genome-wide after Oocyte-NT (Figure 10.4.1). As expected, the RNA-seq libraries of MEF-NT without BrUTP treatment as negative controls are clearly separated from all other libraries with BrUTP treatment regardless of the donor cell types, the origins of oocyte batches and the collection days (Figure 10.4.1). In addition, the libraries of MEF-NT collected at Day 1 and Day 2 after Oocyte-NT are separated from other libraries of mESC-NT and mMyo-NT and are grouped together based on the origins of oocytes (Frog 1 and Frog 2, Figure 10.4.1). Unexpectedly, the RNA-seq libraries of mESC-NT, MEF-NT and mMyo-NT collected at Day 2 after Oocyte-NT are not different from each other for either the difference of cell types or the origins of oocyte batches and it suggests that the reprogrammed nuclei mESC-NT, MEF-NT and mMyo-NT are similar after being reprogrammed by *Xenopus* oocytes for 2 days.

Figure 10.4.1

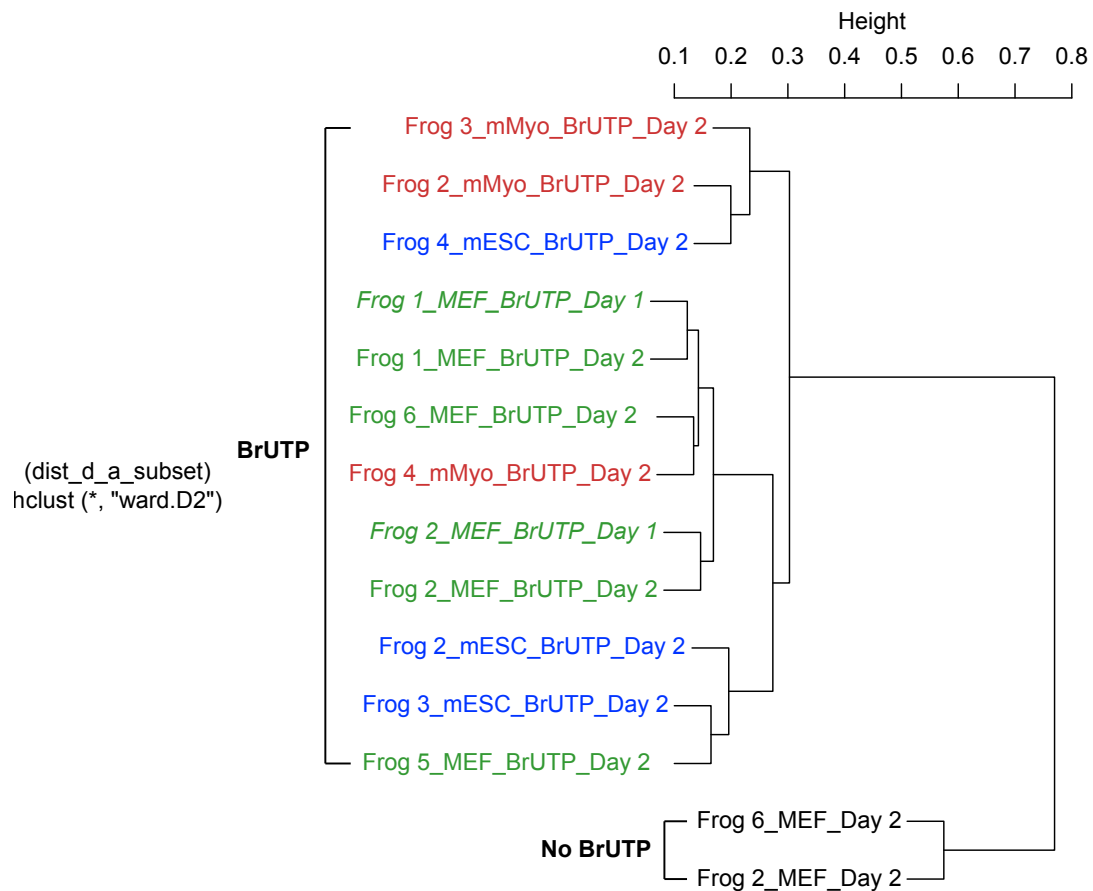


Figure 10.4.1 Reprogrammed transcriptomes of newly synthesized transcripts in mESC, MEF and mMyo were similarly promoted by oocyte factors.

Hierarchical clustering shows that RNA-seq libraries of mESC-NT (in blue), MEF-NT (in green) and mMyo-NT (in red) are grouped by expression of newly synthesized transcripts with or without BrUTP labeling (in black) regardless of cell types and frogs at Day 1 (in *italics*) and Day 2 after Oocyte-NT. Height represents the distance of dissimilarity across samples.

Table 10.4.1

Library	Mapped Reads to Reference Genome			
	<i>Mus musculus</i> (mm10)	<i>Xenopus laevis</i> (xl9.1)		
<i>Frog 1_MEF_BrUTP_Day 1</i>	17,273,206	19,372,578	3nM	#1
<i>Frog 1_MEF_BrUTP_Day 2</i>	19,103,794	17,923,632	3nM	
<i>Frog 2_MEF_BrUTP_Day 1</i>	32,492,438	24,466,766	2.5nM	#2
<i>Frog 2_MEF_Day 2</i>	331,171	5,357,592	0.2nM	
<i>Frog 2_MEF_BrUTP_Day 2</i>	31,605,380	19,553,838	2.5nM	#3
<i>Frog 2_mMyo_BrUTP_Day 2</i>	14,694,612	21,543,039	1.65nM	
<i>Frog 2_mESC_BrUTP_Day 2</i>	32,958,974	23,907,495	1.65nM	#4
<i>Frog 3_mESC_BrUTP_Day 2</i>	46,777,778	18,751,556	1.65nM	
<i>Frog 3_mMyo_BrUTP_Day 2</i>	10,153,737	25,885,533	1.65nM	#5
<i>Frog 4_mMyo_BrUTP_Day 2</i>	16,843,313	18,907,652	2.5nM	
<i>Frog 4_mESC_BrUTP_Day 2</i>	56,573,764	19,038,998	2.5nM	#6
<i>Frog 5_MEF_BrUTP_Day 2</i>	23,075,691	9,662,453	2.5nM	
<i>Frog 6_MEF_Day 2_Library 1*</i>	118,687	1,372,291	0.13nM	#7
<i>Frog 6_MEF_BrUTP_Day 2_Library 1*</i>	11,914,487	7,102,116	2nM	
<i>Frog 6_MEF_Day 2_Library 2*</i>	122,190	1,398,516	0.13nM	#7
<i>Frog 6_MEF_BrUTP_Day 2_Library 2*</i>	12,194,569	7,239,653	2nM	

Table 10.4.1 BrUTP pull-down efficiently enriches mapped reads of newly synthesized transcripts of mESC-NT (in blue), MEF-NT (in green) and mMyo-NT (in red) to mouse genome.

Libraries of MEF-NT at Day 1 after Oocyte-NT are in italics and libraries of MEF-NT without the addition of BrUTP are in black.

Libraries were pooled together with specified concentrations (nM) and sequenced in the same lanes.

* Libraries 1 and 2 of MEF-NT of Frog 6 are technical duplicates and mapped reads are added up together for bioinformatics analysis.

10.5 Appendix V: Gene ontology and KEGG pathway analysis of silenced genes in mESCs, MEFs and mMyos after Oocyte-NT

In mESCs, 3533 genes (Figure 5.4.1.A, page 172) of the original transcriptome are silent after Oocyte-NT, excluding housekeeping genes, are enriched for biological processes regarding development, morphogenesis, cell-cell interaction and signaling transduction pathways via Gene ontology ($p < 0.05$, Figure 10.5.1). Additionally, silenced genes by oocyte factors in mESCs are enriched for pathways, including endocytosis, Rap1 signaling pathway and so on ($p < 0.05$, Figure 10.5.2). Therefore, genes, involved in interaction between cells, accepting extracellular stimuli and signal transduction from extracellular stimuli to intracellular targets, are silenced by oocyte factors.

In MEFs, 2251 silenced genes (Figure 5.4.1.B, page 172) by oocyte factors are enriched for 137 Gene ontology terms of Biological process ($p < 0.05$, Figure 10.5.3). Among these terms, the biological processes selected for functions of MEFs are related to wound healing, extracellular matrix organization, connective tissue development and response to growth factor stimulus (in red, Figure 10.5.3). The pathways come out from the KEGG pathway enrichment of silenced genes in MEFs by oocyte factors, including Rap1 signaling pathway, Axon Guidance, Adherens junction etc ($p < 0.05$, Figure 10.5.4).

In mMyos, 4098 silenced genes (Figure 5.4.1.C, page 172) by oocyte factors are enriched for biological processes via Gene ontology and terms related to functions of mMyos include Striated muscle contraction, Cardiac muscle contraction, Muscle system process, Neuromuscular junction development, Muscle contraction and Heart contraction (in red, $p < 0.05$, Figure 10.5.5). Pathways for mMyos from enrichment of silenced genes include Focal adhesion, Rap1 signaling pathway, Ras signaling pathway etc. ($p < 0.05$, Figure 10.5.6).

Interestingly, some common pathways are seen from enrichment of silenced genes of mESCs, MEFs and mMyos via KEGG pathway, including endocytosis, focal adhesion, axon guidance, adherens junction, Rap1 signaling pathway and Ras signaling pathway (in red, Figure 10.5.2, 10.5.4, 10.5.6). These pathways are not cell-type specific pathways and might be common pathways for cell lines, which have been through long-term cell culture in vitro.

Figure 10.5.1

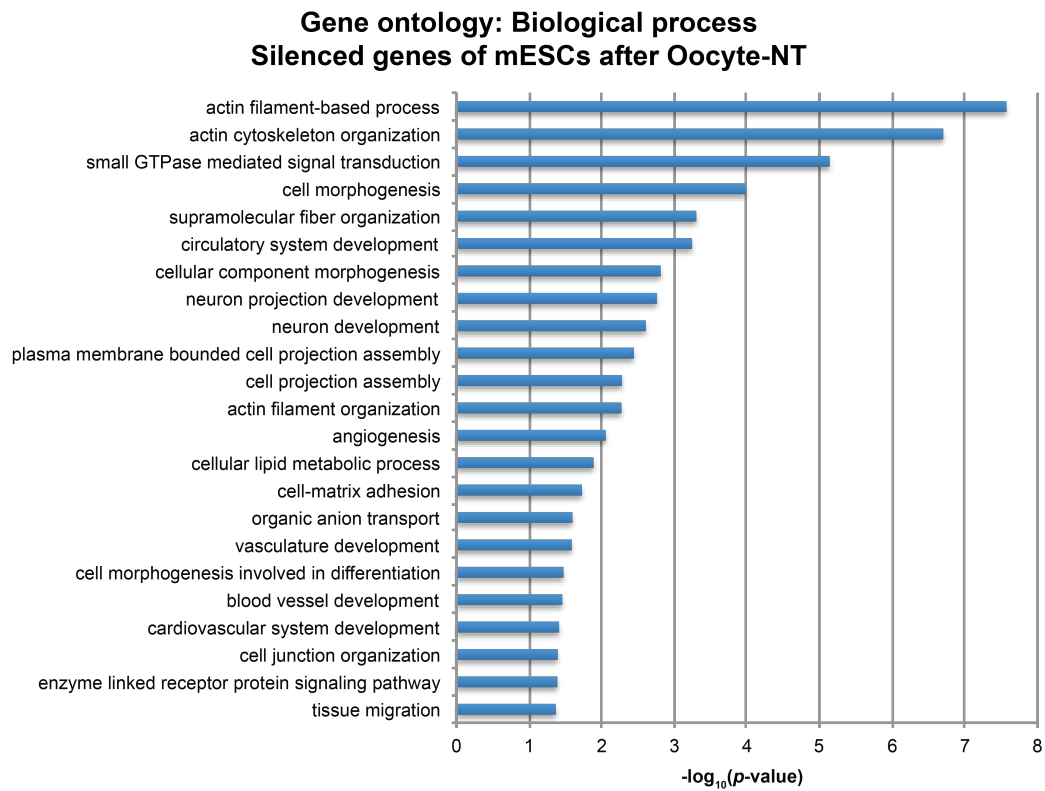


Figure 10.5.1 Silenced genes of mESCs after Oocyte-NT are enriched ($p < 0.05$) for biological processes related to development, morphogenesis, cell-cell interaction and signaling transduction pathways.

Figure 10.5.2

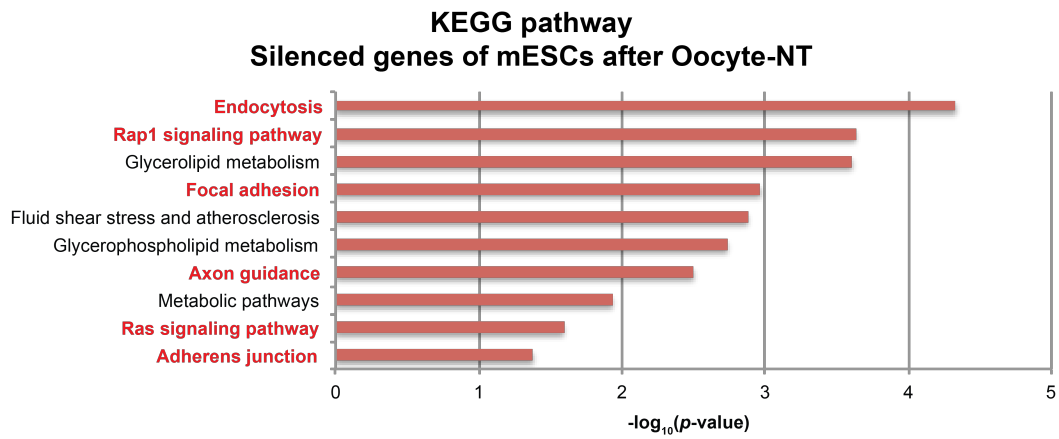


Figure 10.5.2 Silenced genes of mESCs after Oocyte-NT are enriched ($p < 0.05$) for pathways, including endocytosis, Rap1 signaling pathway, glycerolipid metabolism and so on.

Pathways in red are most likely to be common pathways for cell lines.

Figure 10.5.3.A

A.

Gene ontology: Biological process Silenced genes of MEFs after Oocyte-NT (List 1, Rank 1-50)

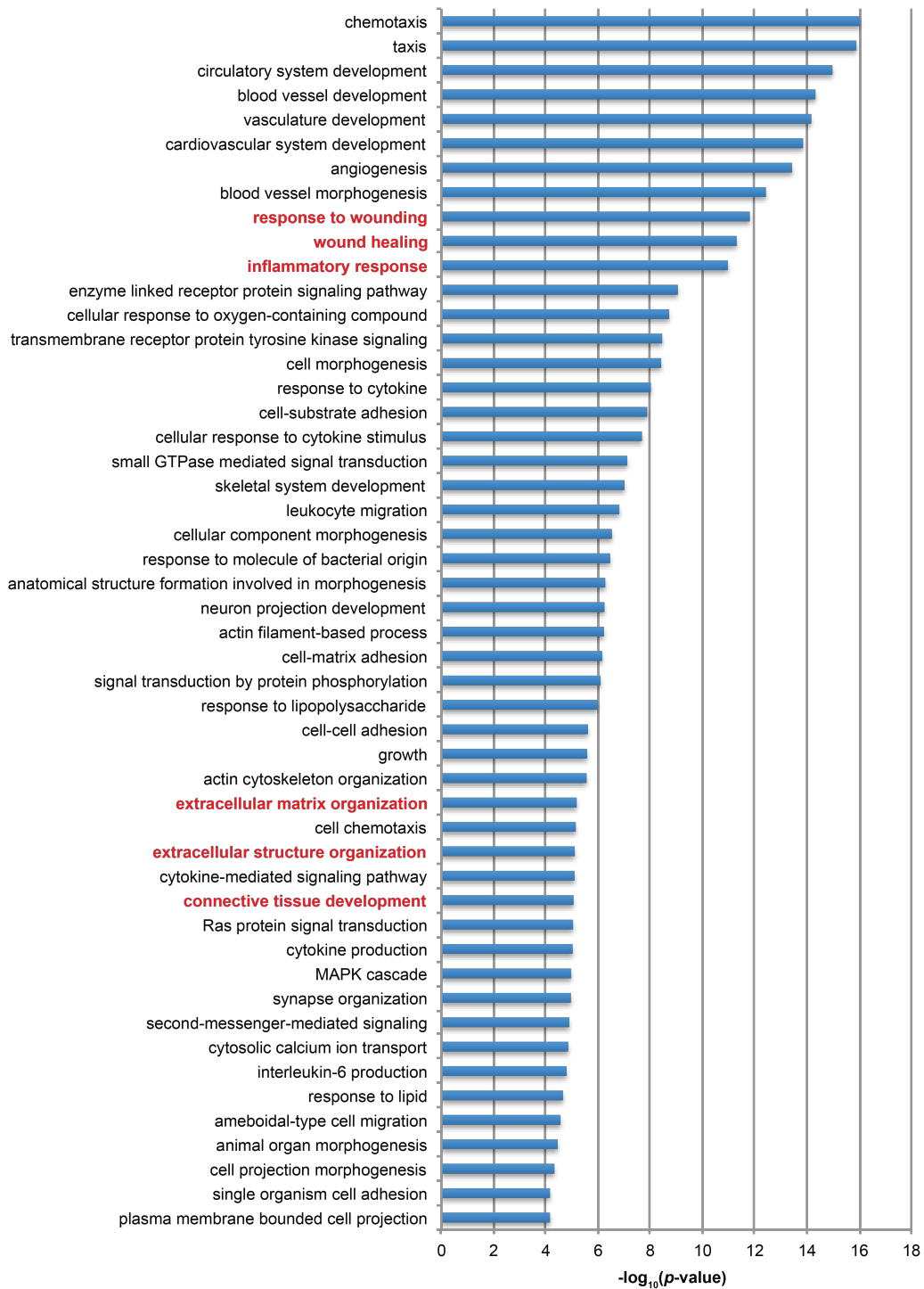


Figure 10.5.3.B

B.

Gene ontology:Biological process
Silenced genes of MEFs after Oocyte-NT (List 2, Rank 51-100)

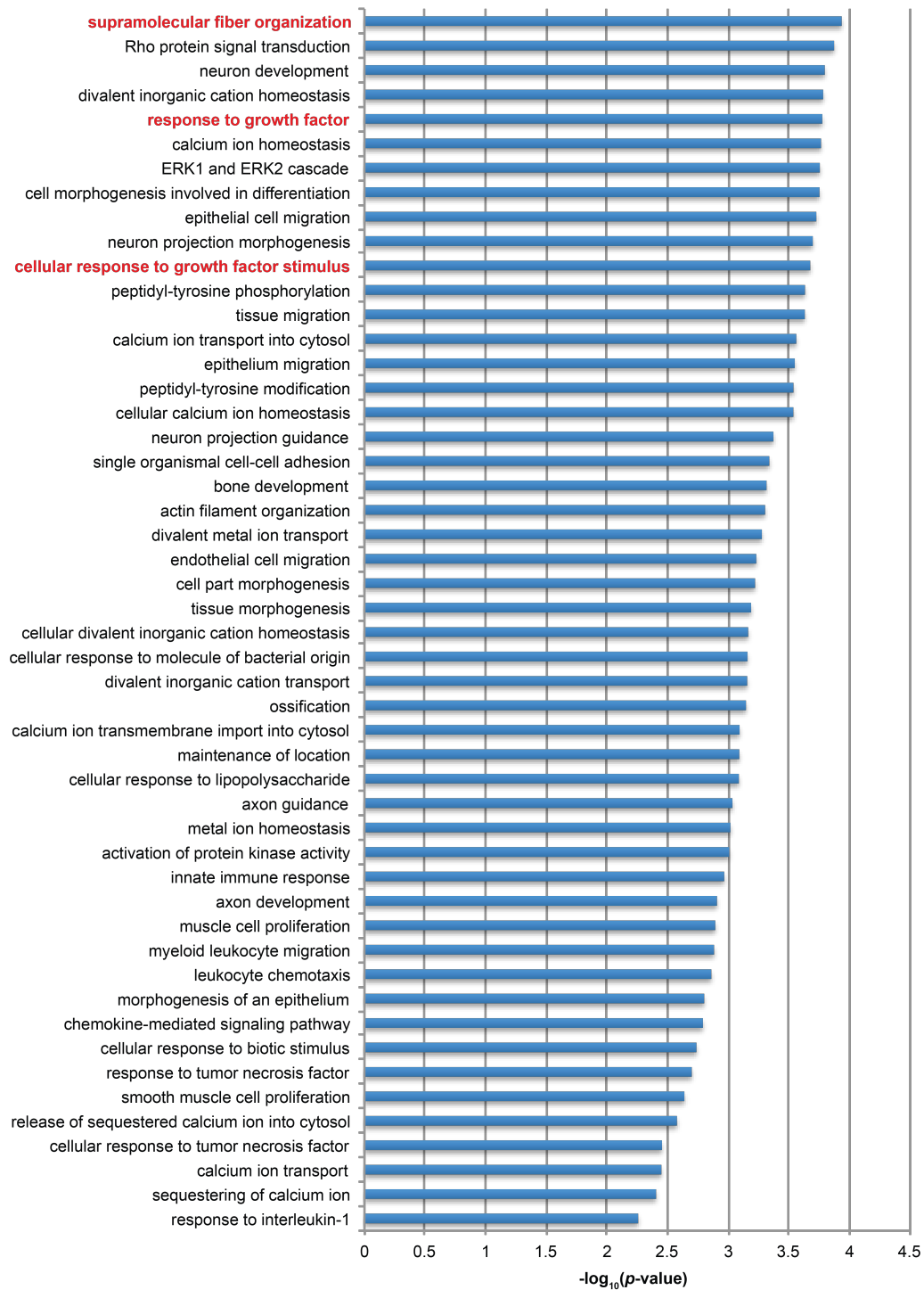


Figure 10.5.3.C

**C. Gene ontology: Biological process
Silenced genes of MEFs after Oocyte-NT (List 3, Rank 101-137)**

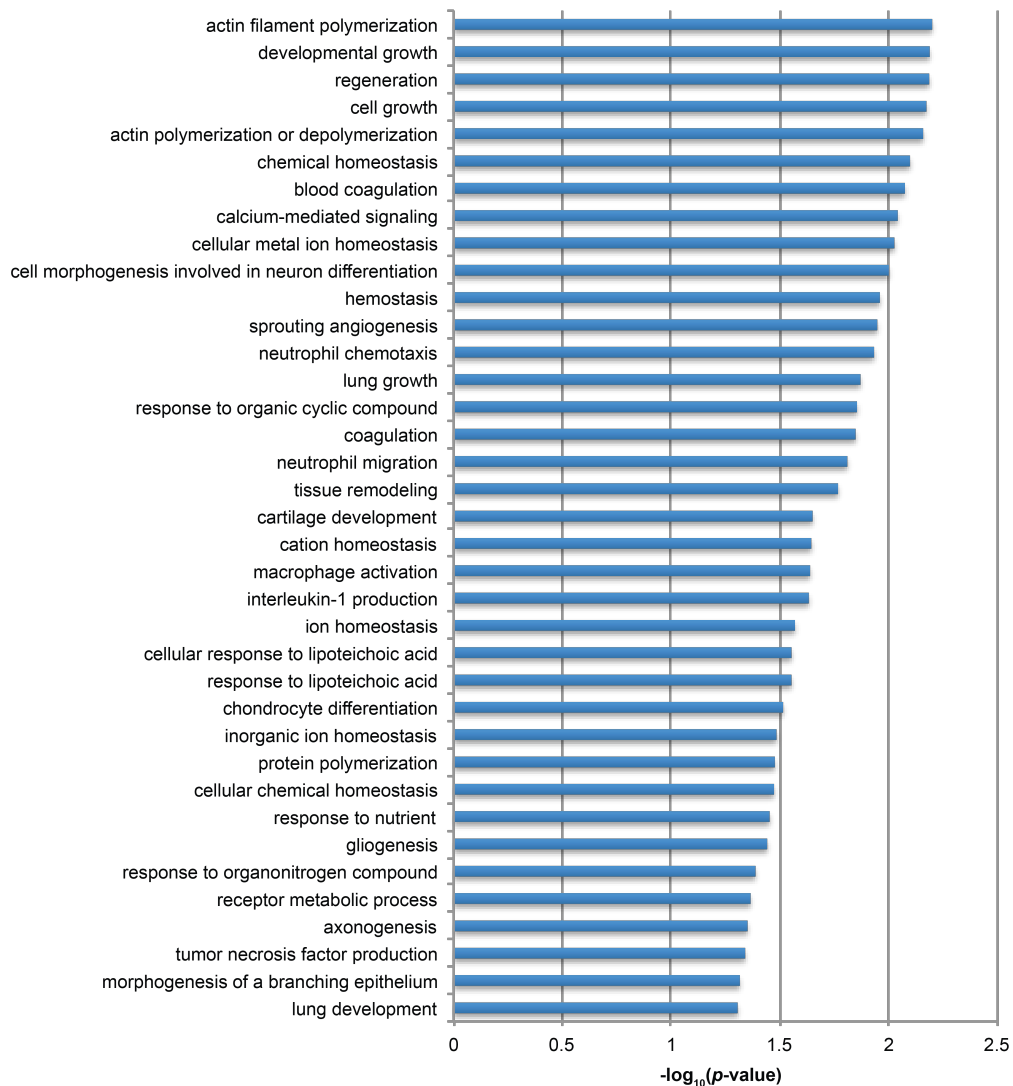


Figure 10.5.3 Silenced genes in MEFs after Oocyte-NT are enriched ($p < 0.05$) for 137 Gene ontology terms of Biological processes and the terms related to the functions of MEFs include response to wounding, wound healing, inflammatory response and so on (in red).

Figure 10.5.4

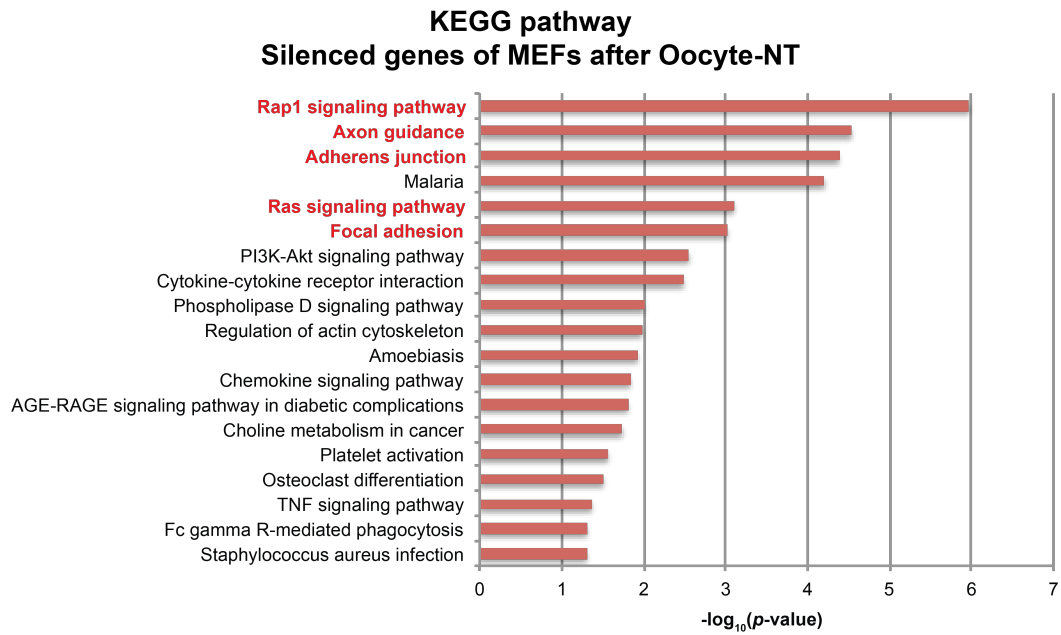


Figure 10.5.4 Silenced genes of MEFs after Oocyte-NT are enriched ($p < 0.05$) for Rap1 signaling pathway, axon guidance, adherens junction and so on.

Pathways in red are most likely to be common pathways for cell lines.

Figure 10.5.5

**Gene ontology: Biological process
Silenced genes of mMyos after Oocyte-NT**

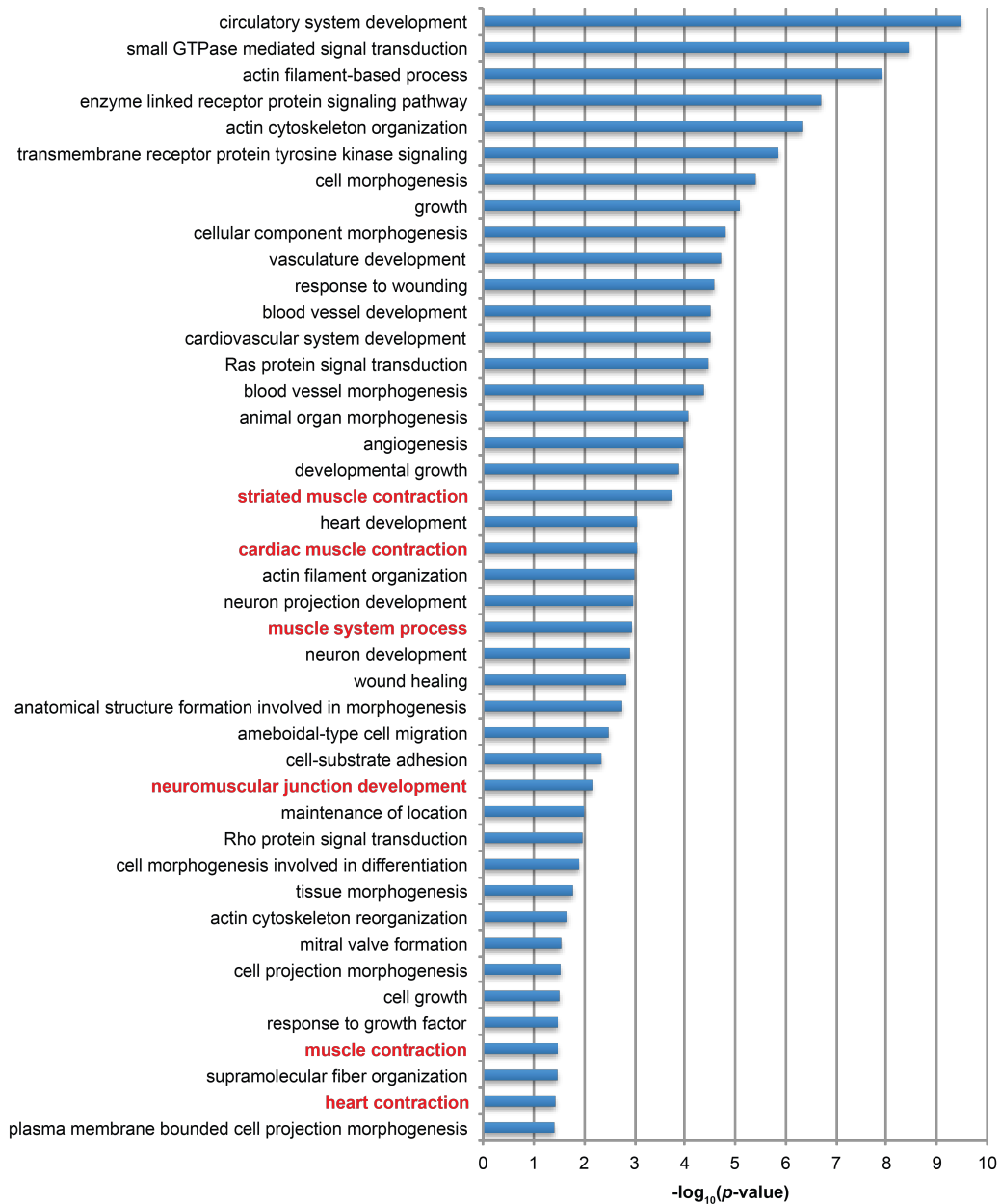


Figure 10.5.5 Silenced genes of mMyos after Oocyte-NT are enriched ($p < 0.05$) for biological processes related to functions of mMyos, such as striated muscle contraction, cardiac muscle contraction, muscle system process, neuromuscular junction development, muscle contraction and heart contraction (in red).

Figure 10.5.6

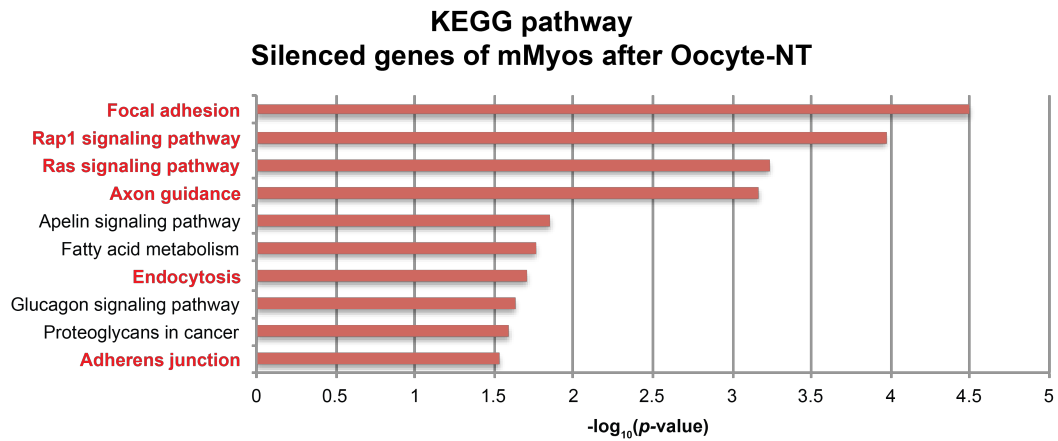


Figure 10.5.6 Silenced genes of mMyos after Oocyte-NT are enriched ($p < 0.05$) for focal adhesion, Rap1 signaling pathway, Ras signaling pathway and so on. Pathways in red are most likely to be common pathways for cell lines.

10.6 Appendix VI: Gene ontology and KEGG pathway analysis of reprogrammable genes in mESCs, MEFs and mMyos after Oocyte-NT

This shows that reprogrammable genes among mESC-NT, MEF-NT and mMyo-NT (FPKM>0 in all three cell types) are involved in many DNA related biological processes, including DNA repair, Cellular response to DNA damage stimulus, DNA replication, DNA duplex unwinding, DNA packaging, DNA conformation change and so on (Figure 10.6.1.A).

Additionally, reprogrammable genes are also involved in many biological processes regarding structural change, modification of chromatin and histone modification (Figure 10.6.1.B and C). For example, reprogrammable genes are involved in Covalent chromatin modification, Chromatin remodeling, Chromatin assembly and disassembly, Chromatin organization, Establishment of protein localization to chromosome and so on (Figure 10.6.1.B). Furthermore, reprogrammable genes are also involved in several histone modification related biological processes, including histone acetylation, deacetylation, methylation, ubiquitination and deubiquitination (Figure 10.6.1.C). Since these chromatin and histone related biological processes could lead to the change of chromatin structure and chromatin accessibility, the oocyte-determined state of reprogrammed transcriptomes have the ability to change the chromatin accessibility and may aid the opening of regulatory elements of resistant genes, which are buried in condensed chromatin structures and silent before Oocyte-NT.

Moreover, the reprogrammable genes are also involved in transcription, translation and embryonic development (Figure 10.6.1.D-F). This means that the reprogrammed transcriptomes have these abilities for the establishment of totipotency, early development of embryos or the needs for SCNR by oocytes.

For the pathways predicted for the functions of reprogrammable genes, some basic functions come out, such as Cell cycle, RNA transport, Spliceosome and so on (in red, Figure 10.6.2.A and B). Interestingly, Oocyte meiosis and Signaling pathways regulating pluripotency of stem cells are also predicted as results (Figure 10.6.2.A). For Oocyte meiosis, this suggests that the functions of *Xenopus* oocytes may also relate to the reprogrammed transcriptomes of donor cells since the oocyte factors perform functions not only for accepting transplanted nuclei but also for maintaining their own cell identity.

Figure 10.6.1.A

A. **Gene ontology: Biological process**
Reprogrammable genes of mESC-NT, MEF-NT and mMyo-NT

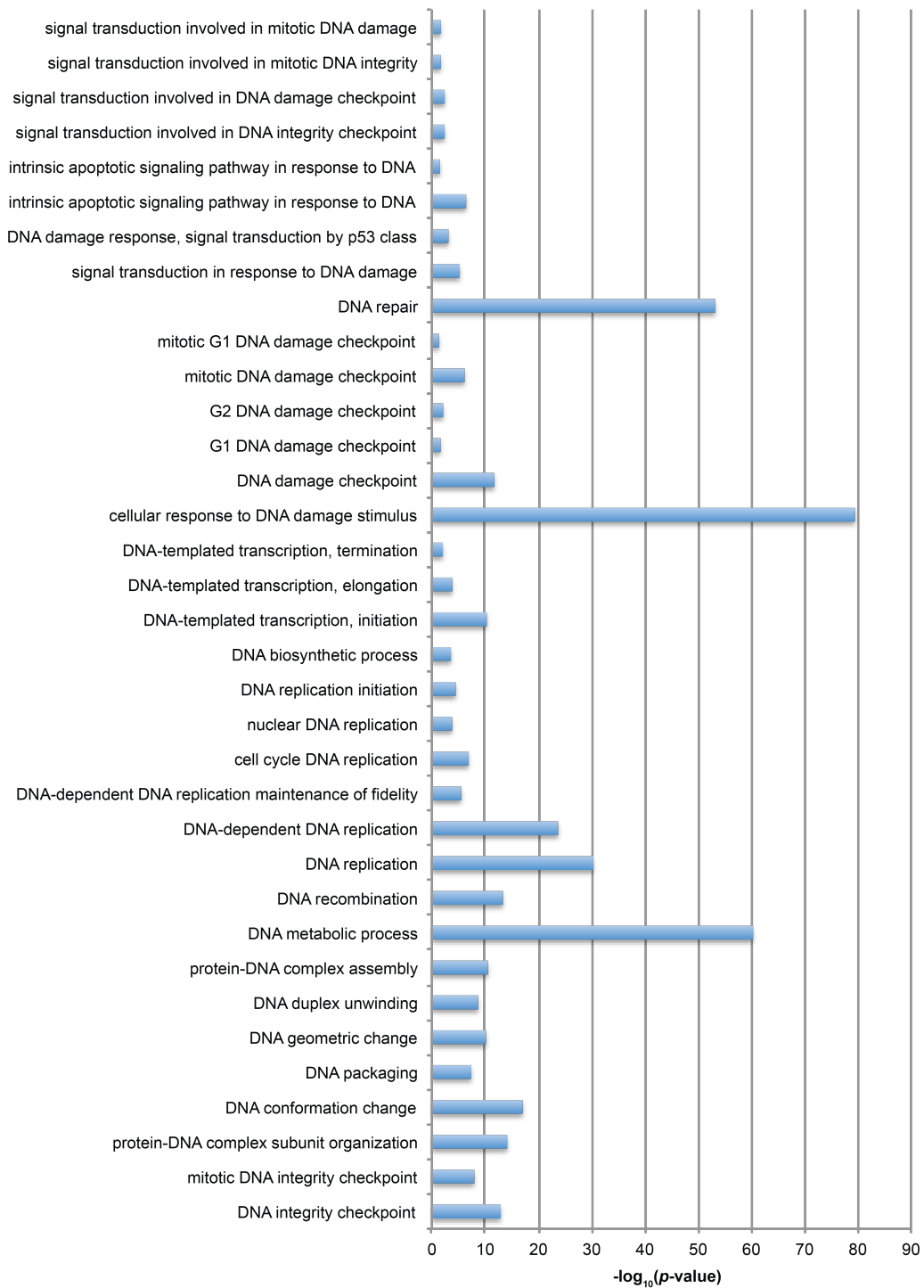
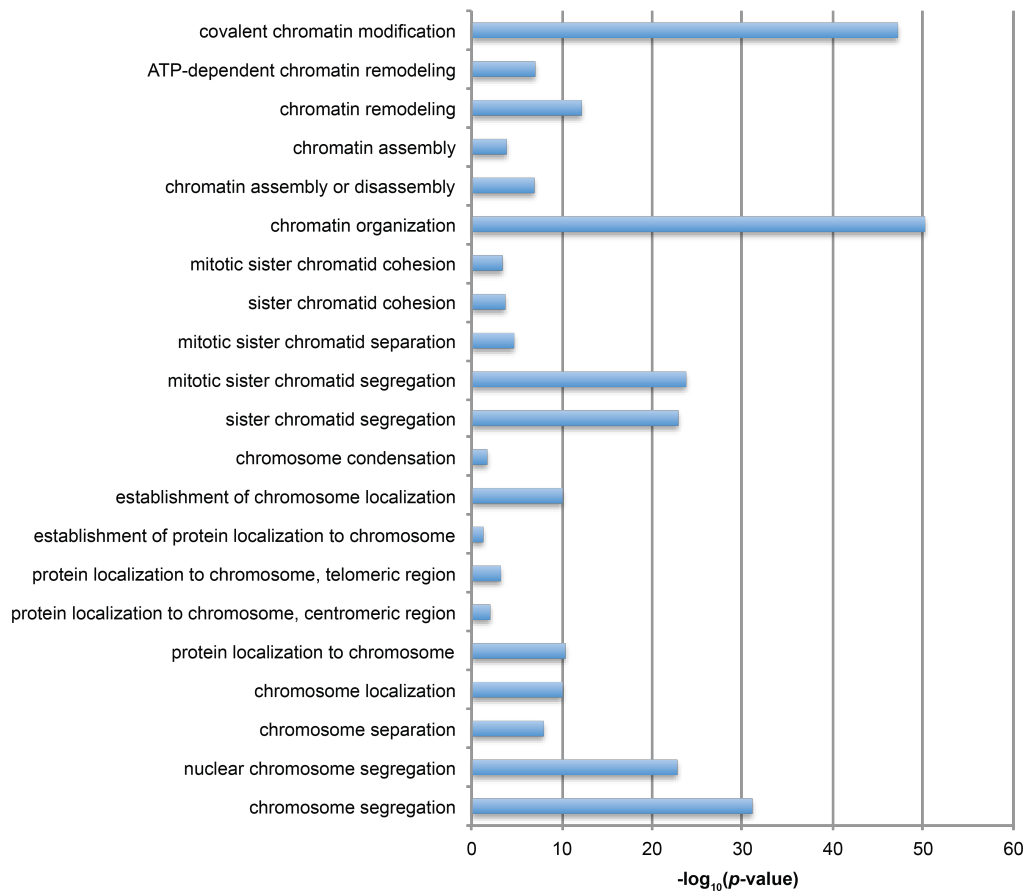


Figure10.6.1.B-C

B.

Gene ontology: Biological process
Reprogrammable genes of mESC-NT, MEF-NT and mMyo-NT



C.

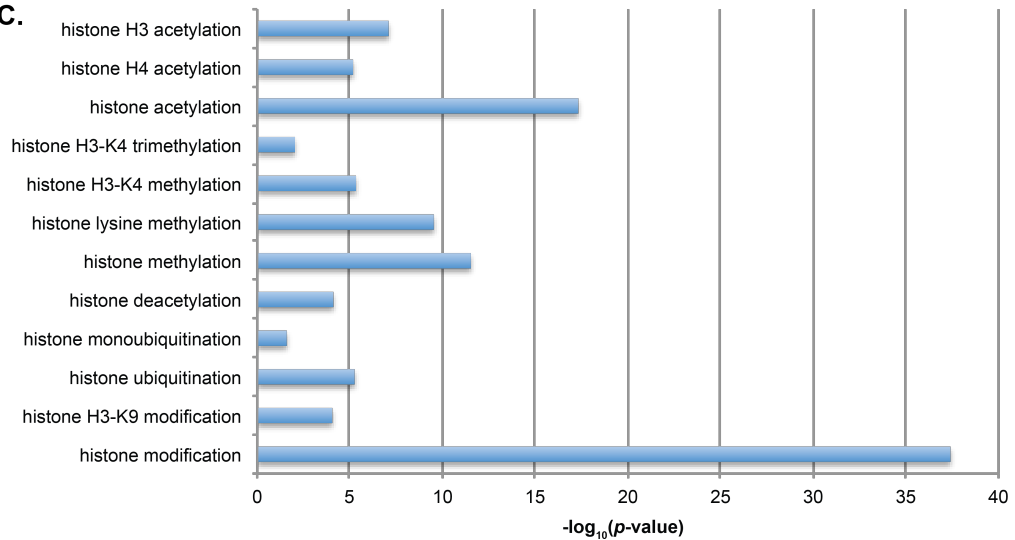


Figure 10.6.1.D-F

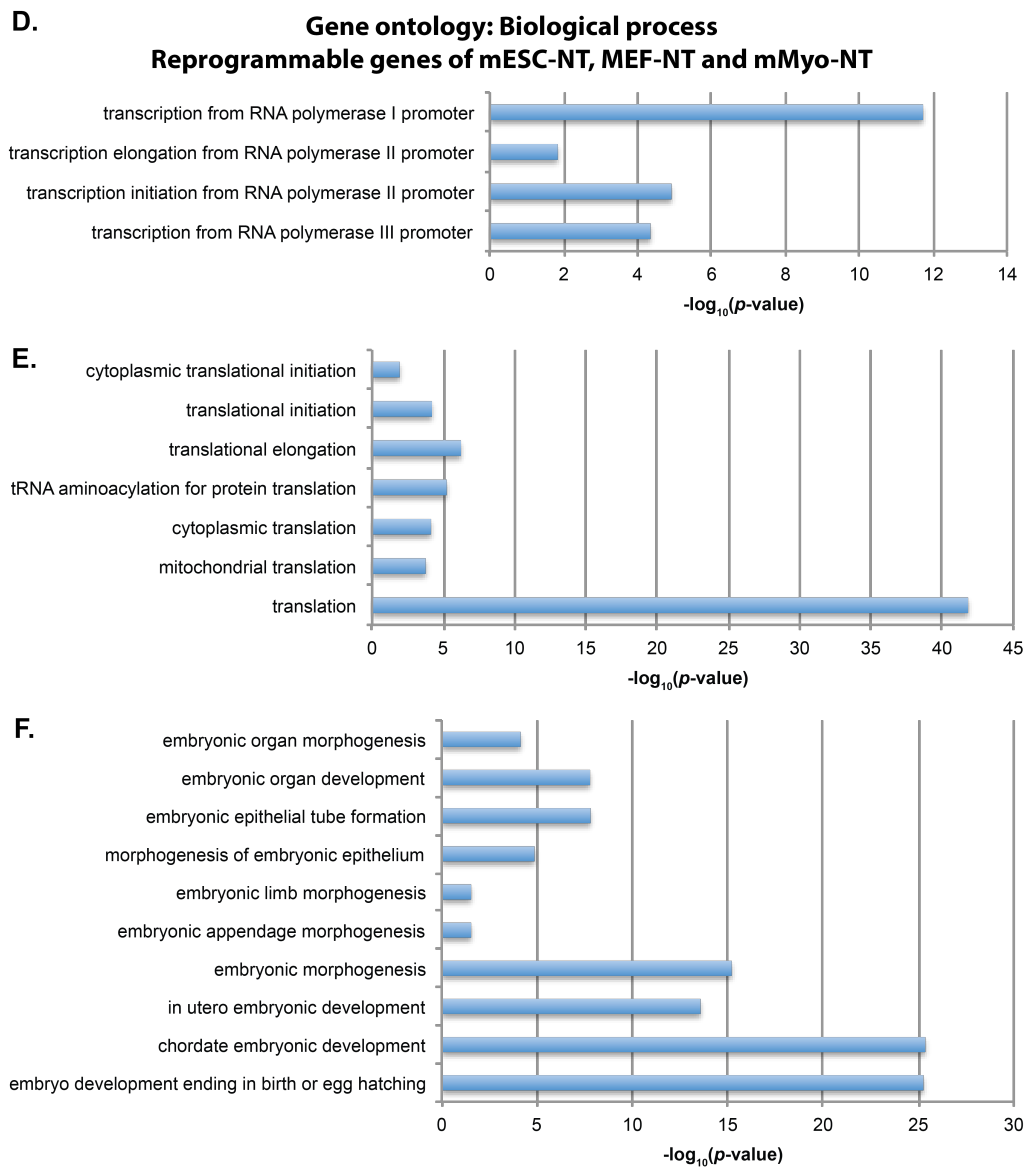


Figure 10.6.1 Reprogrammable genes, which are expressed in mESC-NT, MEF-NT and mMyo-NT, are involved in DNA related biological processes (A), structural change and modification of chromatin (B), histone modification (C), transcription (D), translation (E) and embryonic development (F) ($p < 0.05$).

Figure 10.6.2.A

A.

KEGG pathway
Reprogrammable genes of mESC-NT, MEF-NT and mMyo-NT

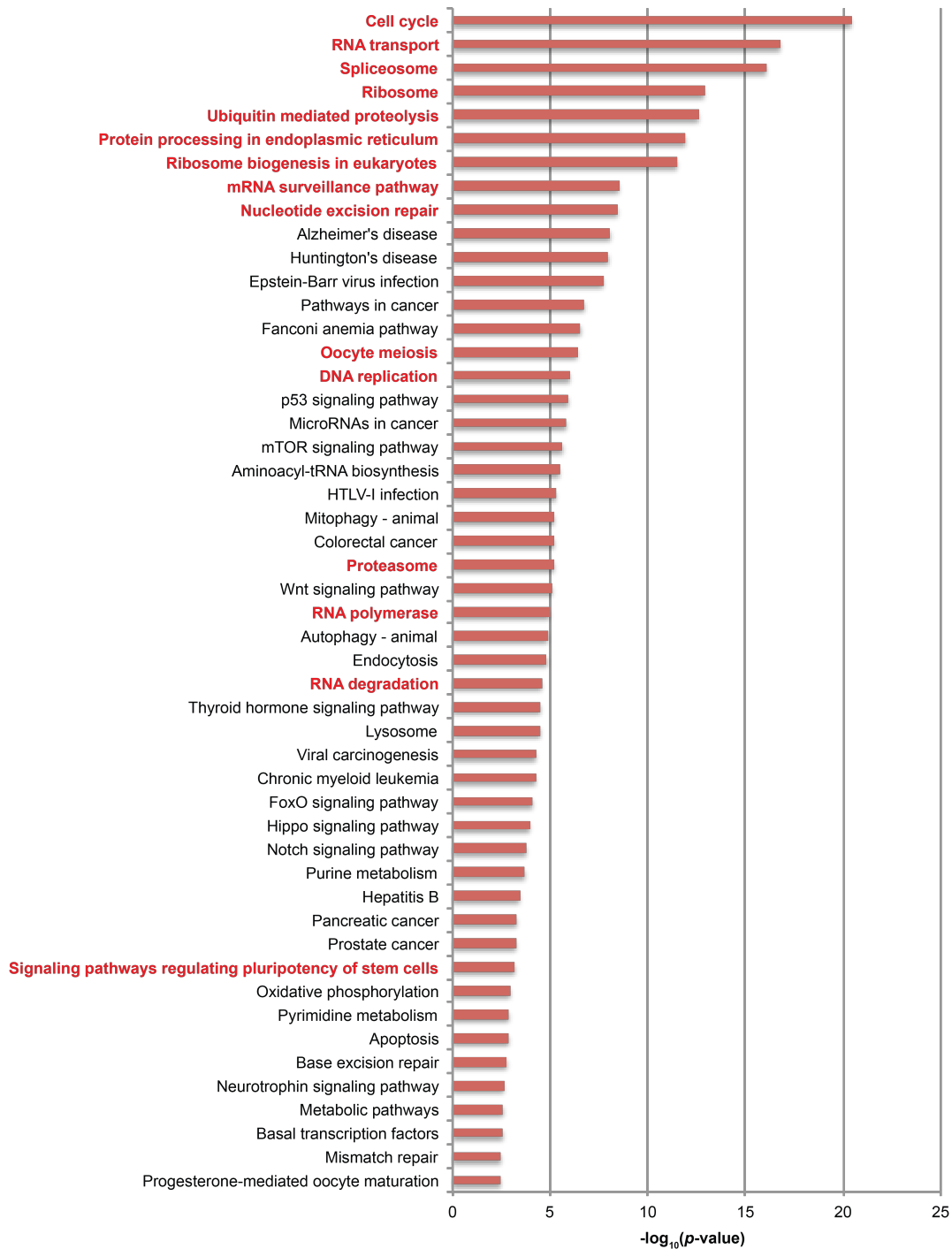


Figure 10.6.2.B

B.

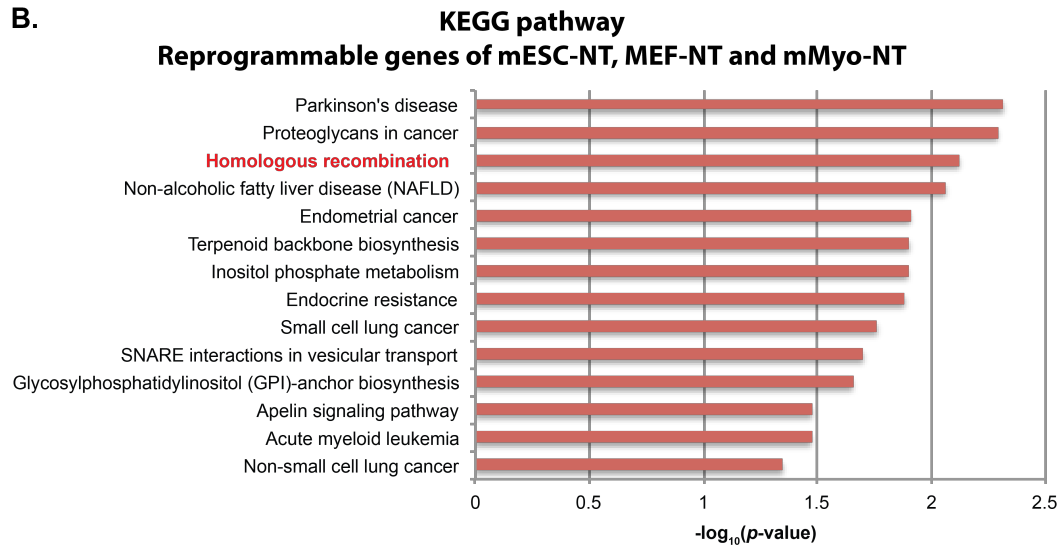


Figure 10.6.2 Reprogrammable genes, which are expressed in mESC-NT, MEF-NT and mMyo-NT, are enriched ($p < 0.05$) for several pathways, including oocyte meiosis and signaling pathways regulating pluripotency of stem cells.

10.7 Appendix VII: Preliminary data validation for reprogrammed transcriptomes of mESCs, MEFs and mMyos in the absence and presence of xklf2-HA overexpression in the Oocyte-NT system

For RNA-seq analysis, there are 2-4 biological replicates for each group (Talbe 10.7.1). Each library was mapped to the mouse genome (*Mus musculus*, mm10) and *Xenopus* genome (*Xenopus laevis*, xla9.1). For mESC-NT, there are an average of 50 million mapped reads to the mouse genome for each library (n=3, Talbe 10.7.1). For MEF-NT, there are an average of 25 million mapped reads to mouse genome for each library (n=2 for Day 1 and n=4 for Day 2, Talbe 10.7.1). For mMyo-NT, there are an average of 15 million reads to mouse genome for each library (n=3, Talbe 10.7.1). Interestingly, the amount of starting material in libraries of mESC-NT and mMyo-NT for sequencing is the same and libraries of mESC-NT and mMyo-NT were pooled together for sequencing in the same sequencing lanes (#3, #4 and #5, Talbe 10.7.1). However, the resulting mapped reads to mouse genome of mESC-NT are 3 times more than mapped reads to mouse genome of mMyo-NT (n=3, Talbe 10.7.1). This suggests that the landscape of chromatin structures of transplanted mESCs may be more open and responsive to transcriptional machineries of oocytes.

Regarding the efficiency of BrUTP pull-down, reads mapped to the mouse genome in MEF-NT with BrUTP addition has 6-9 times more reads than MEF-NT without BrUTP addition (BrUTP/No BrUTP of mapped reads to the mouse genome, Frog 2 and 6, Talbe 10.7.1). Furthermore, reads mapped to the

mouse genome are approximately 1.4 times more than reads mapped to the *Xenopus* genome in MEF-NT with BrUTP addition (mapped reads to the mouse genome/*Xenopus* genome, BrUTP, Frog 2 and 6, Talbe 10.7.1). However, reads mapped to the mouse genome are less than 1/10 of reads mapped to the *Xenopus* genome in MEF-NT without BrUTP addition (mapped reads to the mouse genome/*Xenopus* genome, No BrUTP, Frog 2 and 6, Talbe 10.7.1). Therefore, the newly synthesized transcripts picked up specifically by anti-BrUTP antibodies come mostly from mouse nuclei than *Xenopus* nuclei. In addition, non-specifically bound transcripts in No BrUTP samples after BrUTP pull-down are more from *Xenopus* nuclei than mouse nuclei.

For the time-dependent effect of *xklf2*-HA overexpression, Day 1 and Day 2 samples of MEF-NT of Frog 2 are clustered by the expression level of genes genome wide in respect of *xklf2*-HA treatment, BrUTP addition, cell type difference and batch effect of oocytes (Day 1 and Day 2, Frog 2, Figure 10.7.1). Yet, Day 1 and Day 2 samples of MEF-NT of Frog 1 are clustered by the expression level of genes genome wide based on *xklf2*-HA treatment but not cell type difference and batch effect of oocytes (Day 1 and Day 2, Frog 1, Figure 10.7.1). It suggests that the genome-wide difference may not be significant between Day 1 and Day 2 samples of MEF-NT and mMyo-NT under *xklf2*-HA overexpression.

For the cell-type specific effects of *xklf2*-HA overexpression, Day 2 samples of mESC-NT, MEF-NT and mMyo-NT are clustered by the expression level of

genes genome wide only based on *xklf2*-HA treatment (Figure 10.7.1). Interestingly, cell type differences and batch effects of oocytes do not impact on the clustering of Day 2 samples of mESC-NT, MEF-NT and mMyo-NT (Figure 10.7.1). This resembles the clustering results in Chapter 5 that reprogrammed transcriptomes are very similar at Day 2 after Oocyte-NT regardless of cell type difference and batch effects of oocytes (Figure 10.4.1, page 329). It indicates that mESC-NT, MEF-NT and mMyo-NT under the treatment of *xklf2*-HA overexpression are similar to each other at Day 2 after Oocyte-NT.

Figure 10.7.1

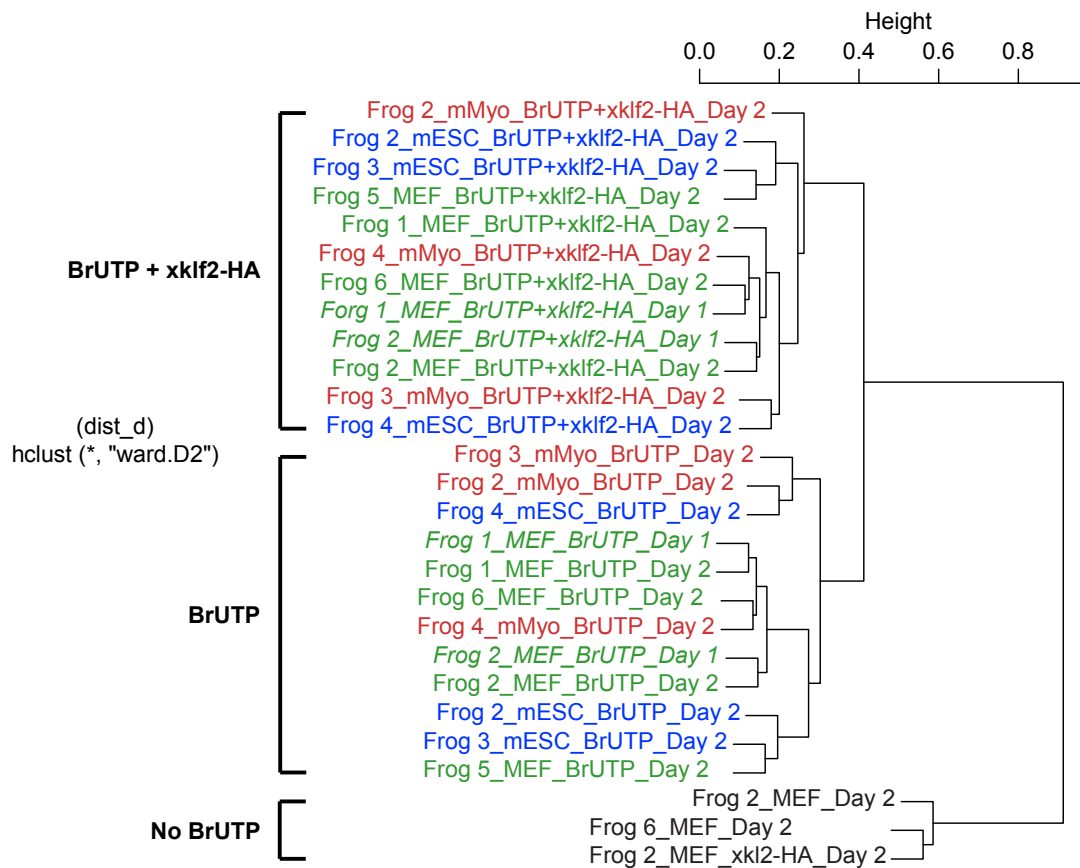


Figure 10.7.1 Hierarchical clustering shows that RNA-seq libraries of mESC-NT (in blue), MEF-NT (in green) and mMyo-NT (in red) are grouped by expression of newly synthesized transcripts with or without BrUTP labeling (in black) in the presence or absence of xklf2-HA overexpression regardless of cell types and frogs at Day 1 (in *italics*) and Day 2 after Oocyte-NT.

Height represents the distance of dissimilarity across samples.

Table 10.7.1

Library	Mapped Reads to Reference Genome		Input	
	<i>Mus musculus</i> (mm10)	<i>Xenopus laevis</i> (xl9.1)		
<i>Frog 1_MEF_BrUTP_Day 1</i>	17,273,206	19,372,578	3 nM each	#1
<i>Frog 1_MEF_BrUTP+xklf2-HA_Day 1</i>	22,931,297	19,110,888		
Frog 1_MEF_BrUTP_Day 2	19,103,794	17,923,632		
Frog 1_MEF_BrUTP+xklf2-HA_Day 2	17,430,993	20,019,654		
<i>Frog 2_MEF_BrUTP_Day 1</i>	32,492,438	24,466,766	2.5 nM	#2
<i>Frog 2_MEF_BrUTP+xklf2-HA_Day 1</i>	31,289,535	31,719,250	2.5 nM	
Frog 2_MEF_Day 2	331,171	5,357,592	0.2 nM	
<i>Frog 2_MEF_BrUTP_Day 2</i>	31,605,380	19,553,838	2.5 nM	
Frog 2_MEF_xklf2-HA_Day 2	221,033	4,965,468	0.2 nM	
<i>Frog 2_MEF_BrUTP+xklf2-HA_Day 2</i>	33,349,772	17,195,567	2.5 nM	
<i>Frog 2_mMyo_BrUTP_Day 2</i>	14,694,612	21,543,039	1.65 nM each	#3
<i>Frog 2_mMyo_BrUTP+xklf2-HA_Day 2</i>	11,762,475	20,986,915		
<i>Frog 2_mESC_BrUTP_Day 2</i>	32,958,974	23,907,495		
<i>Frog 2_mESC_BrUTP+xklf2-HA_Day 2</i>	32,494,714	23,422,173		
<i>Frog 3_mESC_BrUTP_Day 2</i>	46,777,778	18,751,556	1.65 nM each	#4
<i>Frog 3_mESC_BrUTP+xklf2-HA_Day 2</i>	62,890,625	17,062,270		
<i>Frog 3_mMyo_BrUTP_Day 2</i>	10,153,737	25,885,533		
<i>Frog 3_mMyo_BrUTP+xklf2-HA_Day 2</i>	20,617,526	17,200,578		
<i>Frog 4_mMyo_BrUTP_Day 2</i>	16,843,313	18,907,652	2.5 nM each	#5
<i>Frog 4_mMyo_BrUTP+xklf2-HA_Day 2</i>	14,066,332	15,045,898		
<i>Frog 4_mESC_BrUTP_Day 2</i>	56,573,764	19,038,998		
<i>Frog 4_mESC_BrUTP+xklf2-HA_Day 2</i>	45,804,067	17,651,509		
<i>Frog 5_MEF_BrUTP_Day 2</i>	23,075,691	9,662,453		
<i>Frog 5_MEF_BrUTP+xklf2-HA_Day 2</i>	29,437,910	11,161,692		
Frog 6_MEF_Day 2_Library 1*	118,687	1,372,291	0.13 nM	#6
<i>Frog 6_MEF_BrUTP_Day 2_Library 1*</i>	11,914,487	7,102,116	2 nM	
<i>Frog 6_MEF_BrUTP+xklf2-HA_Day 2_Library 1*</i>	10,225,459	7,959,804	2 nM	
Frog 6_MEF_Day 2_Library 2*	122,190	1,398,516	0.13 nM	#7
<i>Frog 6_MEF_BrUTP_Day 2_Library 2*</i>	12,194,569	7,239,653	2 nM	
<i>Frog 6_MEF_BrUTP+xklf2-HA_Day 2_Library 2*</i>	10,401,693	8,089,930	2 nM	

Table 10.7.1 BrUTP pull-down efficiently enriches mapped reads of newly synthesized transcripts of mESC-NT (in blue), MEF-NT (in green) and mMyo-NT (in red) to mouse genome.

Libraries of MEF-NT at Day 1 after Oocyte-NT are in italics and libraries of MEF-NT without BrUTP addition are in black.

Libraries were pooled together with specified concentrations and sequenced in the same lanes.

* Libraries 1 and 2 of MEF-NT of Frog 6 are technical duplicates and mapped reads are added up together for bioinformatics analysis.

10.8 Appendix VIII: The involvement of xklf2-HA downstream genes in gene regulation during SCNR by oocytes

To investigate the regulation of downstream genes by xklf2-HA overexpression, genes regulated by xklf2-HA overexpression at Day 1 after Oocyte-NT and genes regulated by xklf2-HA overexpression from Day 1 to Day 2 after Oocyte-NT are compared (Figure 10.8.1 and 10.8.2).

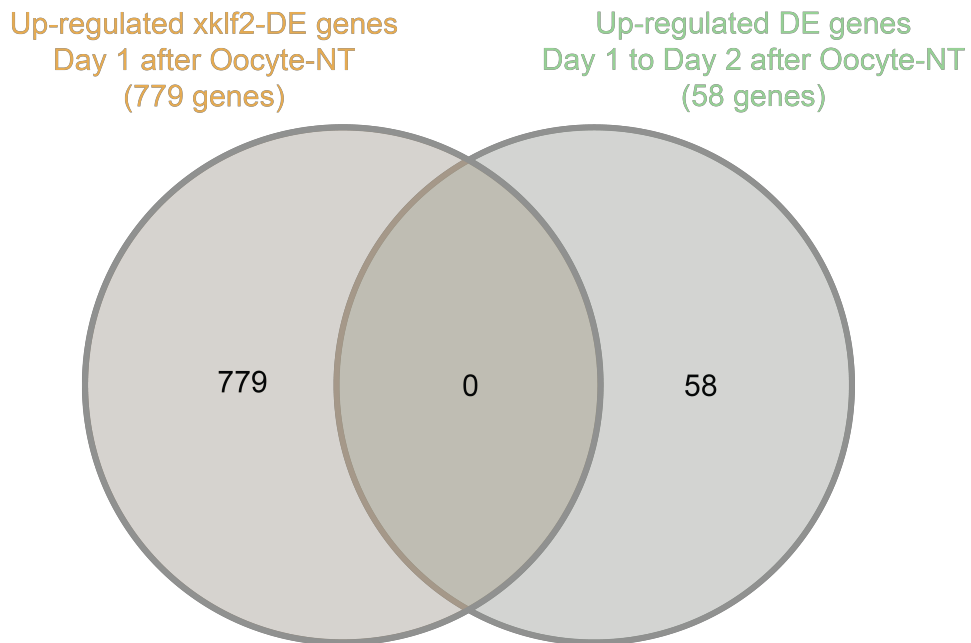
For the DE genes regulated by xklf2-HA overexpression, xklf2-DE genes at Day 1 are regulated by xklf2-HA overexpression by comparing Control groups and xklf2-HA groups at Day 1 after Oocyte-NT (FDR<0.1, Figure 6.2.1, page 196); DE genes from Day 1 to Day 2 after Oocyte-NT are regulated by xklf2-HA overexpression by comparing xklf2-HA groups at Day 1 and xklf2-HA groups at Day 2 after Oocyte-NT (FDR<0.1, Figure 6.2.2, page 200). Since there are no common genes between xklf2-DE at Day 1 and DE genes from Day 1 to Day 2 after Oocyte-NT regulated by xklf2-HA overexpression, it shows that no genes are up-regulated ($\log_2FC > 0$, FDR<0.1) or down-regulated ($\log_2FC < 0$, FDR<0.1) by xklf2-HA overexpression continuously and significantly from time points before Day 1 after Oocyte-NT to time points after Day 1 after Oocyte-NT (Figure 10.8.1). Furthermore, it shows that downstream genes of xklf2-HA regulate different sets of DE genes after Day 1 after Oocyte-NT because DE genes up-regulated and down-regulated by xklf2-HA overexpression before and after Day 1 after Oocyte-NT are totally different (Figure 10.8.1).

While observing genes regulated by xklf2-HA overexpression regardless of FDR, 693 genes are up-regulated ($\log_2FC > 1$) and 351 genes are down-regulated ($\log_2FC < 1$) by xklf2-HA overexpression continuously from time points before Day 1 after Oocyte-NT to time points after Day 1 after Oocyte-NT (Figure 10.8.2).

Nevertheless, the majority of genes regulated by xklf2-HA overexpression are different before and after Day 1 after Oocyte-NT. The difference of temporal gene regulation suggests that downstream genes of xklf2-HA may be involved in gene regulation by xklf2-HA overexpression or the epigenetic barriers of donor cell nuclei are removed temporally and xklf2-HA targets genes are regulated at different time points as a result.

Figure 10.8.1

A. Up-regulated xklf2-DE genes and DE genes by xklf2-HA overexpression ($\log_2FC > 0$, $FDR < 0.1$)



B. Down-regulated xklf2-DE genes and DE genes by xklf2-HA overexpression ($\log_2FC < 0$, $FDR < 0.1$)

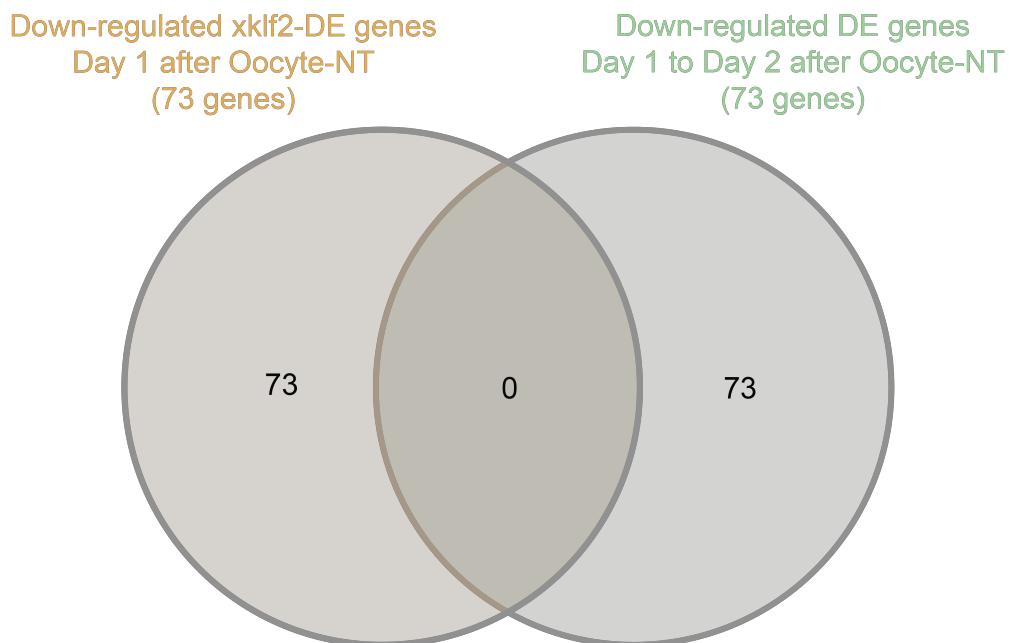
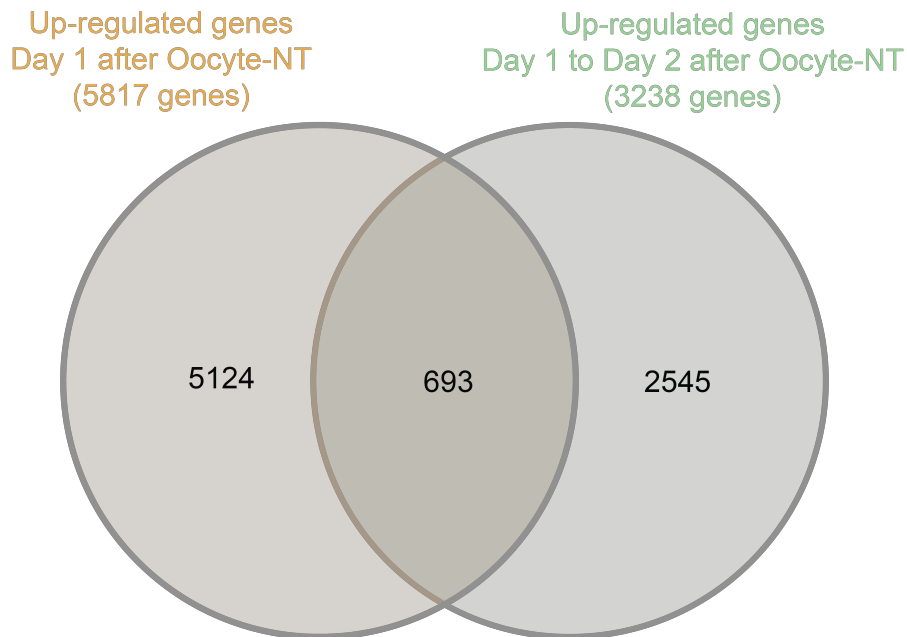


Figure 10.8.1 xklf2-HA overexpression regulates different sets of genes between Day 1 after Oocyte-NT and Day 1 to Day 2 after Oocyte-NT ($FDR < 0.1$).

Up-regulated (A) and down-regulated (B) xklf2-DE genes by xklf2-HA overexpression at Day 1 after Oocyte-NT are compared with up-regulated and down-regulated DE genes by xklf2-HA overexpression from Day 1 to Day 2 after Oocyte-NT, respectively.

Figure 10.8.2

A. Up-regulated genes by xklf2-HA overexpression ($\log_2FC > 1$)



B. Down-regulated genes by xklf2-HA overexpression ($\log_2FC < -1$)

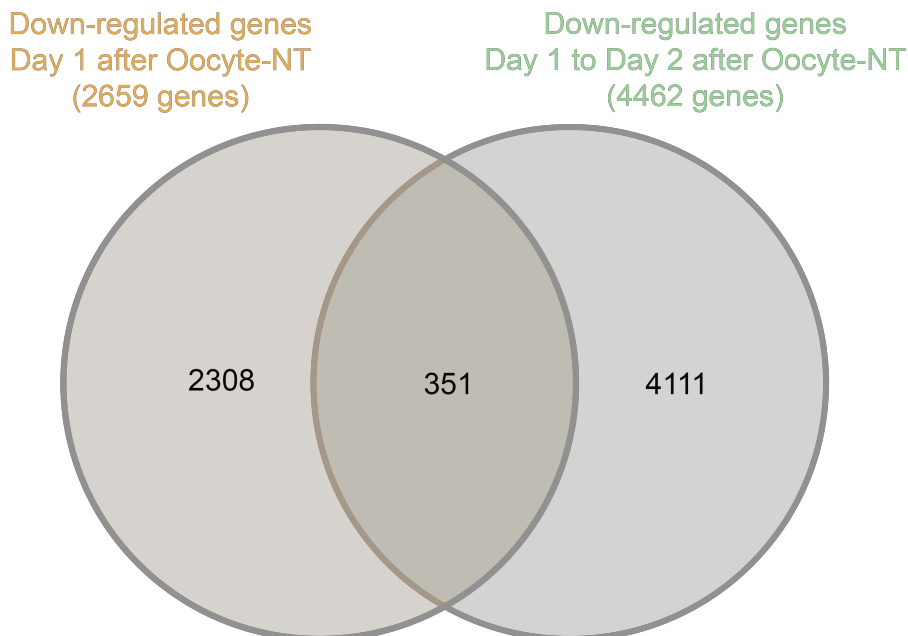


Figure 10.8.2 xklf2-HA overexpression regulates different sets of genes between Day 1 after Oocyte-NT and Day 1 to Day 2 after Oocyte-NT with some shared genes regardless of FDR.

Up-regulated (A) and down-regulated (B) genes by xklf2-HA overexpression at Day 1 after Oocyte-NT are compared with up-regulated and down-regulated genes from Day 1 to Day 2 after Oocyte-NT, respectively.

10.9 Appendix IX: xklf2-HA overexpression activates genes at nil/low expression level and represses genes at high expression level during SCNR by *Xenopus* oocytes

To look deeper into the effect of xklf2-HA overexpression on gene expression (\log_2FC (xklf2-HA/Control)) while expression of genes in mESCs, MEFs and mMyos is reprogrammed by oocyte factors to an oocyte-steady level from Day 0 to Day 2 after Oocyte-NT, xklf2-DE genes are grouped by the expression level of newly synthesized transcripts in reprogrammed transcriptomes of mESCs, MEFs and mMyos in the absence of xklf2-HA (FPKM of Control groups) at Day 2 after Oocyte-NT (Figure 10.9.1, 10.9.2 and 10.9.3; p -value, Table 10.9.1, 10.9.2 and 10.9.3).

Additionally, the number of xklf2-DE genes at different expression level in control groups shows that xklf2-HA overexpression tends to up-regulate genes at nil (FPKM=0) or low (FPKM<1) expression level in control groups and down-regulate genes at high expression level (FPKM>4) in control groups (Figure 10.9.1, 10.9.2 and 10.9.3; p -value, Table 10.9.1, 10.9.2 and 10.9.3). In mESC-NT, MEF-NT and mMyo-NT, more than half of up-regulated xklf2-DE genes are at nil/low expression level ($0 \leq FPKM < 1$) and are up-regulated by xklf2-HA overexpression by 8-fold to 256-fold ($3 < \text{mean of } \log_2FC < 8$, Figure 10.9.1.A, 10.9.2.A, 10.9.3.A).

More than 80% down-regulated xklf2-DE genes are at high expression level in control groups (FPKM>4) and are down-regulated by xklf2-HA overexpression by more than 2-fold to 32-fold ($1 < \text{mean of } \log_2FC < 5$, Figure 10.9.1.B, 10.9.2.B

and 10.9.3.B). Therefore, xklf2-HA overexpression are more effective to activate genes at nil/low expression level in control groups and repress genes at high expression level in control groups when reprogrammed transcriptomes of donor cells in control groups are at oocyte-steady state.

Since expression level of genes correlates with chromatin accessibility and the amount of factors that participate in transcription, the gene activation by xklf2-HA overexpression suggests that xklf2-HA overexpression can access regulatory elements reside in closed chromatin when the linked genes are resist to be activated by oocyte factors.

Figure 10.9.1

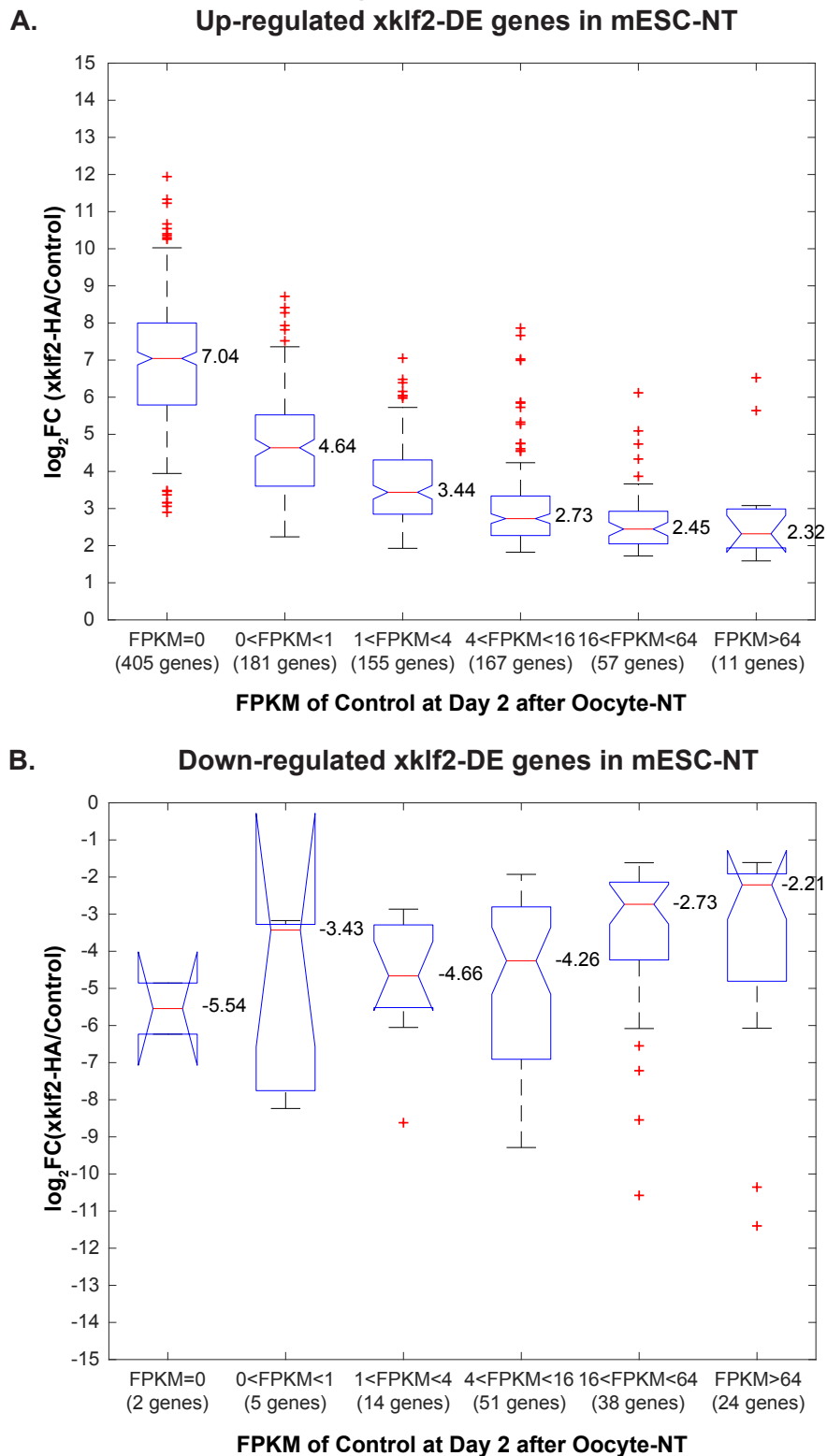


Figure 10.9.2

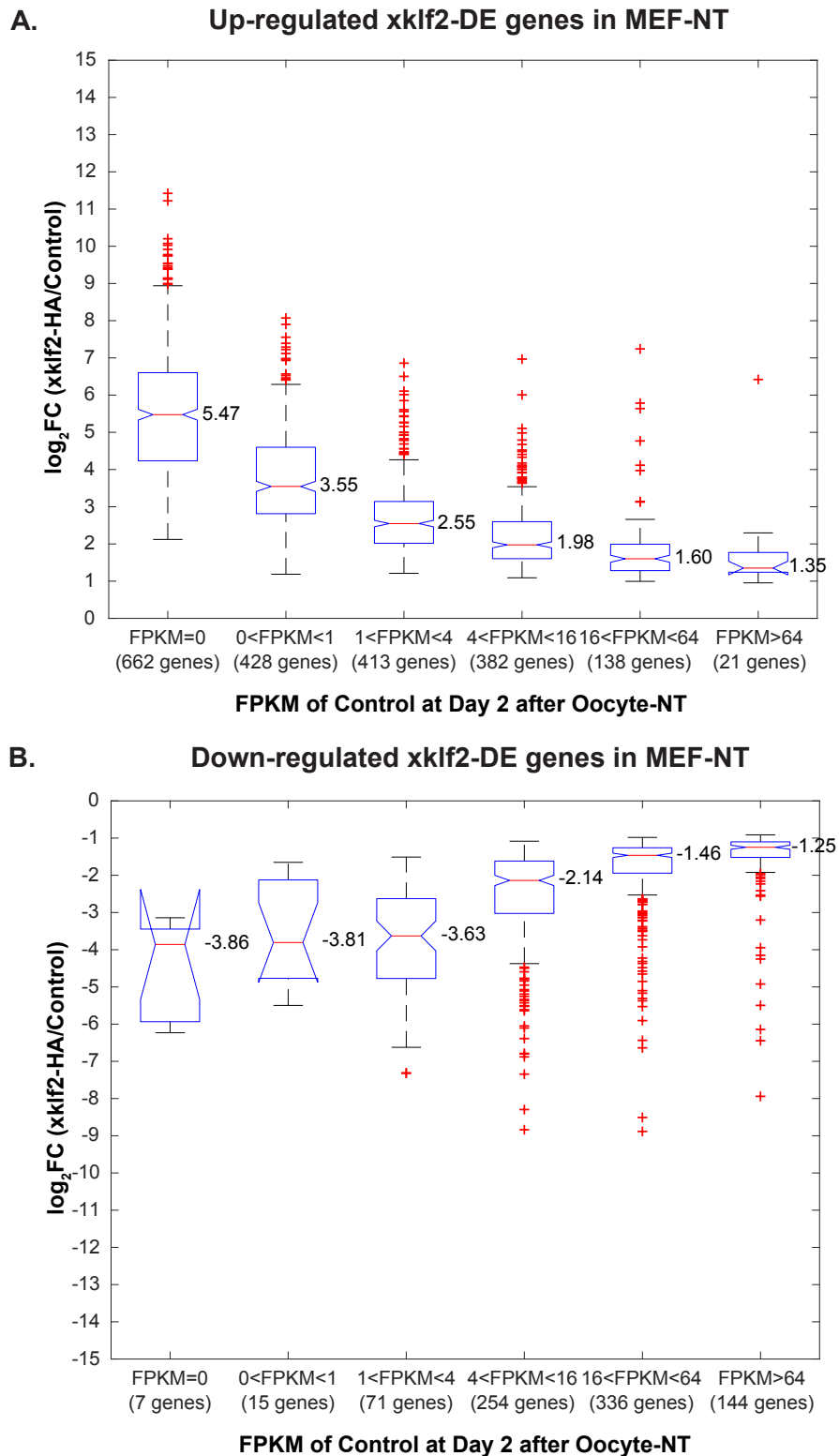


Figure 10.9.2 xklf2-HA overexpression up-regulate xklf2-DE genes at lower expression level with higher \log_2FC and down-regulate xklf2-DE genes at higher expression with lower \log_2FC in MEF-NT at Day 2 after Oocyte-NT significantly (p -value is in Table 10.9.2). Most up-regulated xklf2-DE genes are at nil or low expression level and most down-regulated xklf2-DE genes are at high expression level.

Figure 10.9.3

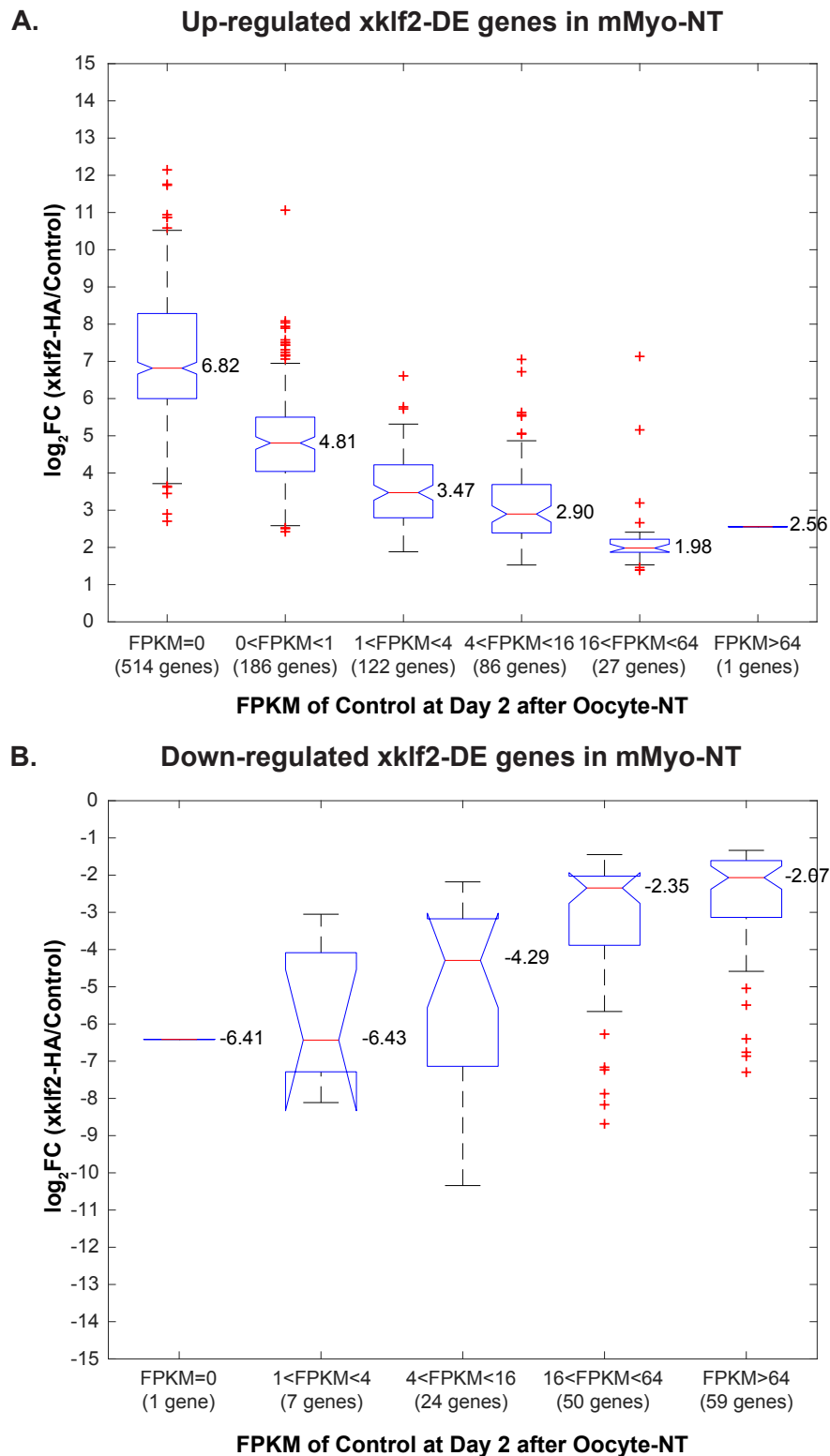


Table 10.9.1

A. Up-regulated xklf2-DE genes in mESC-NT

		log ₂ FC (xklf2-HA/Control)						p-value
		FPKM=0	0<FPKM<1	1<FPKM<4	4<FPKM<16	16<FPKM<64	FPKM>64	
log ₂ FC (xklf2-HA/Control)	FPKM=0							
	0<FPKM<1	2.19E-41						
	1<FPKM<4	1.29E-61	8.56E-14					
	4<FPKM<16	1.25E-70	2.17E-32	1.20E-09				
	16<FPKM<64	1.03E-32	7.02E-21	5.85E-10	0.012			
	FPKM>64	5.32E-07	1.85E-04	0.0094	0.1899	0.7517		

B. Down-regulated xklf2-DE genes in mESC-NT

		log ₂ FC (xklf2-HA/Control)						p-value
		FPKM=0	0<FPKM<1	1<FPKM<4	4<FPKM<16	16<FPKM<64	FPKM>64	
log ₂ FC (xklf2-HA/Control)	FPKM=0							
	0<FPKM<1	0.8571						
	1<FPKM<4	0.3333	0.8932					
	4<FPKM<16	0.5596	0.7959	0.9174				
	16<FPKM<64	0.1001	0.0661	0.0049	9.48E-04			
	FPKM>64	0.1358	0.069	0.0148	8.93E-04	0.2506		

Table 10.9.1 p-value for Figure 10.9.1, p-value<0.05 is in red

Table 10.9.2

A. Up-regulated xklf2-DE genes in MEF-NT

		log ₂ FC (xklf2-HA/Control)						p-value
		FPKM=0	0<FPKM<1	1<FPKM<4	4<FPKM<16	16<FPKM<64	FPKM>64	
log ₂ FC (xklf2-HA/Control)	FPKM=0							
	0<FPKM<1	4.19E-65						
	1<FPKM<4	2.04E-131	2.88E-39					
	4<FPKM<16	8.33E-144	8.28E-76	1.84E-20				
	16<FPKM<64	2.93E-70	1.69E-54	4.56E-33	1.63E-11			
	FPKM>64	3.42E-13	4.86E-12	1.14E-09	8.07E-06	0.0869		

B. Down-regulated xklf2-DE genes in MEF-NT

		log ₂ FC (xklf2-HA/Control)						p-value
		FPKM=0	0<FPKM<1	1<FPKM<4	4<FPKM<16	16<FPKM<64	FPKM>64	
log ₂ FC (xklf2-HA/Control)	FPKM=0							
	0<FPKM<1	2.05E-01						
	1<FPKM<4	1.24E-01	7.94E-01					
	4<FPKM<16	6.66E-04	3.60E-03	4.76E-11				
	16<FPKM<64	3.62E-05	4.25E-07	5.71E-26	1.83E-21			
	FPKM>64	4.42E-05	3.31E-08	2.54E-25	1.34E-29	1.47E-09		

Table 10.9.2 p-value for Figure 10.9.2, p-value<0.05 is in red

Table 10.9.3

A. Up-regulated xklf2-DE genes in mMyo-NT

		log ₂ FC (xklf2-HA/Control)						p-value
		FPKM=0	0<FPKM<1	1<FPKM<4	4<FPKM<16	16<FPKM<64	FPKM>64	
log ₂ FC (xklf2-HA/Control)	FPKM=0							
	0<FPKM<1	1.00E-46						
	1<FPKM<4	4.93E-60	2.04E-20					
	4<FPKM<16	8.94E-45	7.60E-21	0.0027				
	16<FPKM<64	7.79E-17	1.67E-13	1.04E-09	1.04E-06			
	FPKM>64	0.0845	0.0973	0.2659	0.5637	0.2652		

B. Down-regulated xklf2-DE genes in mMyo-NT

		log ₂ FC (xklf2-HA/Control)						p-value
		FPKM=0	0<FPKM<1	1<FPKM<4	4<FPKM<16	16<FPKM<64	FPKM>64	
log ₂ FC (xklf2-HA/Control)	FPKM=0							
	0<FPKM<1							
	1<FPKM<4	1						
	4<FPKM<16	0.5326		0.4084				
	16<FPKM<64	0.1852		2.70E-03	1.42E-04			
	FPKM>64	0.1333		5.06E-04	8.80E-07	0.0197		

Table 10.9.3 p-value for Figure 10.9.3, p-value<0.05 is in red

10.10 Appendix X: xklf2-HA overexpression down-regulates xklf2-DE genes to xklf2-oocyte level at Day 2 after Oocyte-NT

To look further into down-regulation of xklf2-DE genes by xklf2-HA overexpression in different cell nuclei at Day 2 after Oocyte-NT, expression level of xklf2-DE genes in mESC-NT, MEF-NT and mMyo-NT are compared and \log_2FC of xklf2-DE genes between Control groups and xklf2-HA groups in mESC-NT, MEF-NT and mMyo-NT are compared (Figure 10.10.1, 10.10.2 and 10.10.3).

For 12 xklf2-DE genes down-regulated in all 3 cell types at Day 2 after Oocyte-NT, when expression level (FPKM) is higher, the \log_2FC (xklf2-HA/Control) is lower (Figure 10.10.1). It indicates there is a fixed xklf2-oocyte level for each gene being down-regulated by xklf2-HA overexpression. Likewise, this down-regulation of genes by xklf2-HA overexpression to xklf2-oocyte level is also observed for xklf2-DE down-regulated in only 1 cell type by xklf2-HA overexpression (Figure 10.10.3). However, this down-regulation of genes by xklf2-HA overexpression is not obvious for down-regulated xklf2-DE in 2 cell types possibly due to that the number of reference genes is not sufficient for comparison (xklf2-DE genes < 40, Figure 10.10.2).

Figure 10.10.1

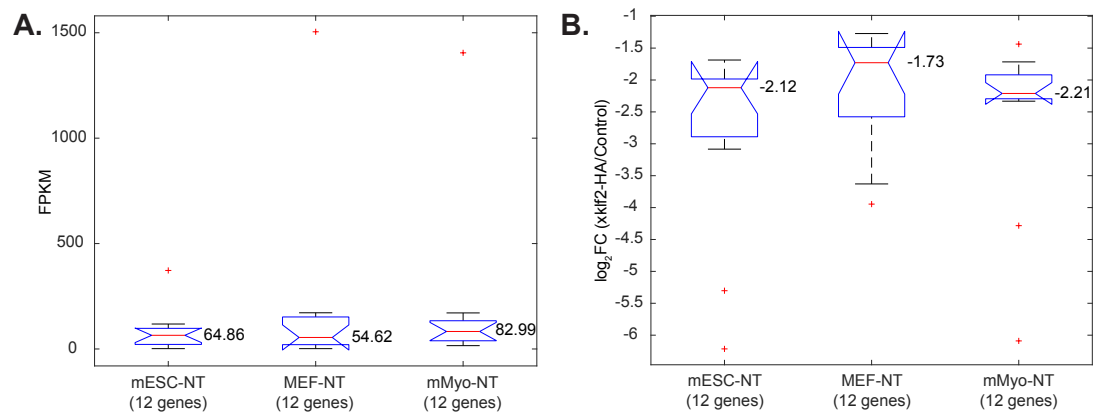


Figure 10.10.1 Expression level (FPKM) in the absence of xklf2-HA and \log_2FC of shared down-regulated xklf2-DE genes in all 3 cell types.

Figure 10.10.2

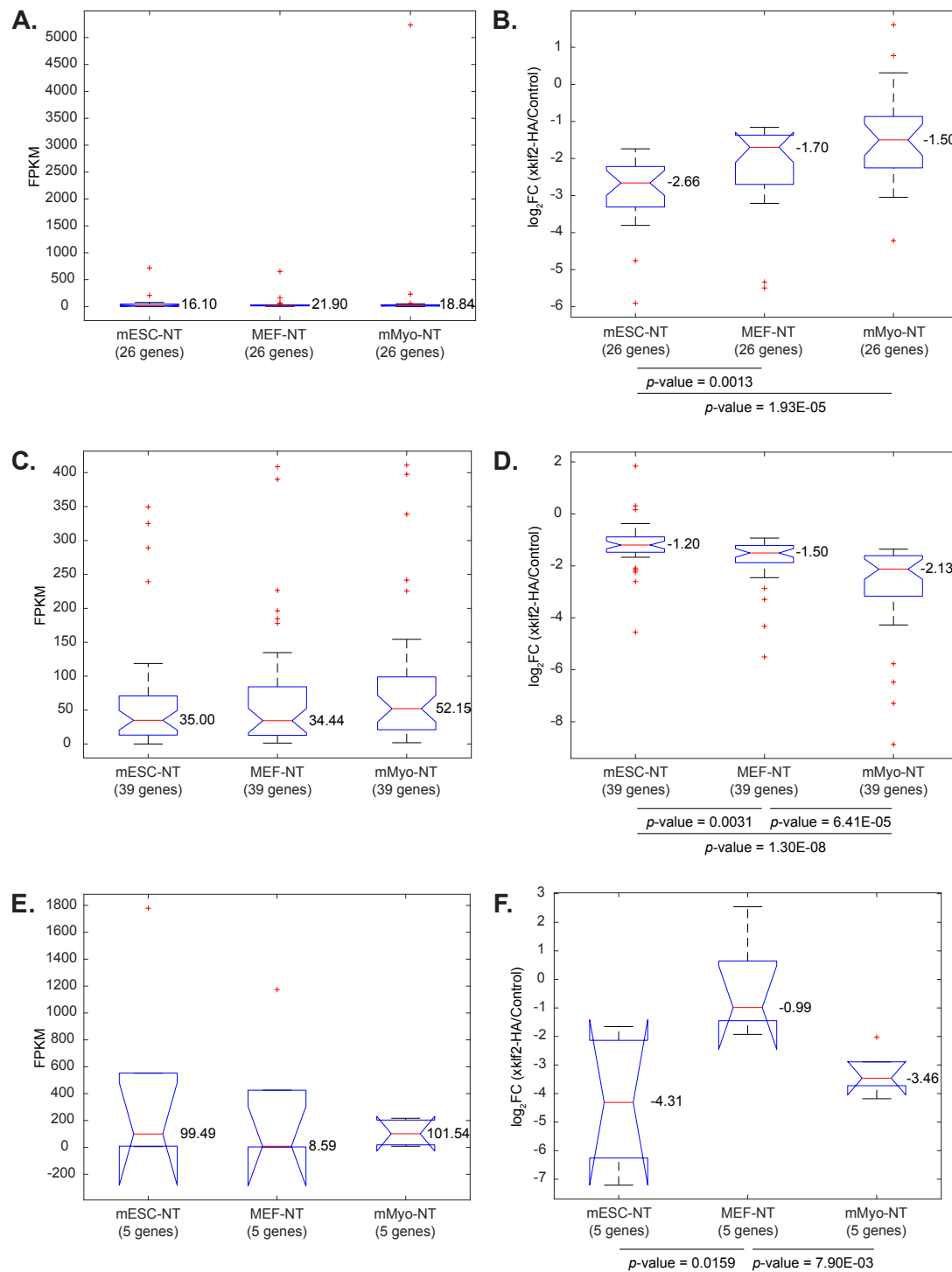


Figure 10.10.2 Shared down-regulated xklf2-DE genes in 2 cell types have higher expression level (FPKM) in the absence of xklf2-HA and down-regulated by xklf2-HA overexpression with lower \log_2FC than the 3rd cell type at Day 2 after Oocyte-NT.

(A, B) Down-regulated xklf2-DE genes shared between mESC-NT and MEF-NT

(C, D) Down-regulated xklf2-DE genes shared between MEF-NT and mMyo-NT

(E, F) Down-regulated xklf2-DE genes shared between mESC-NT and mMyo-NT

Figure 10.10.3

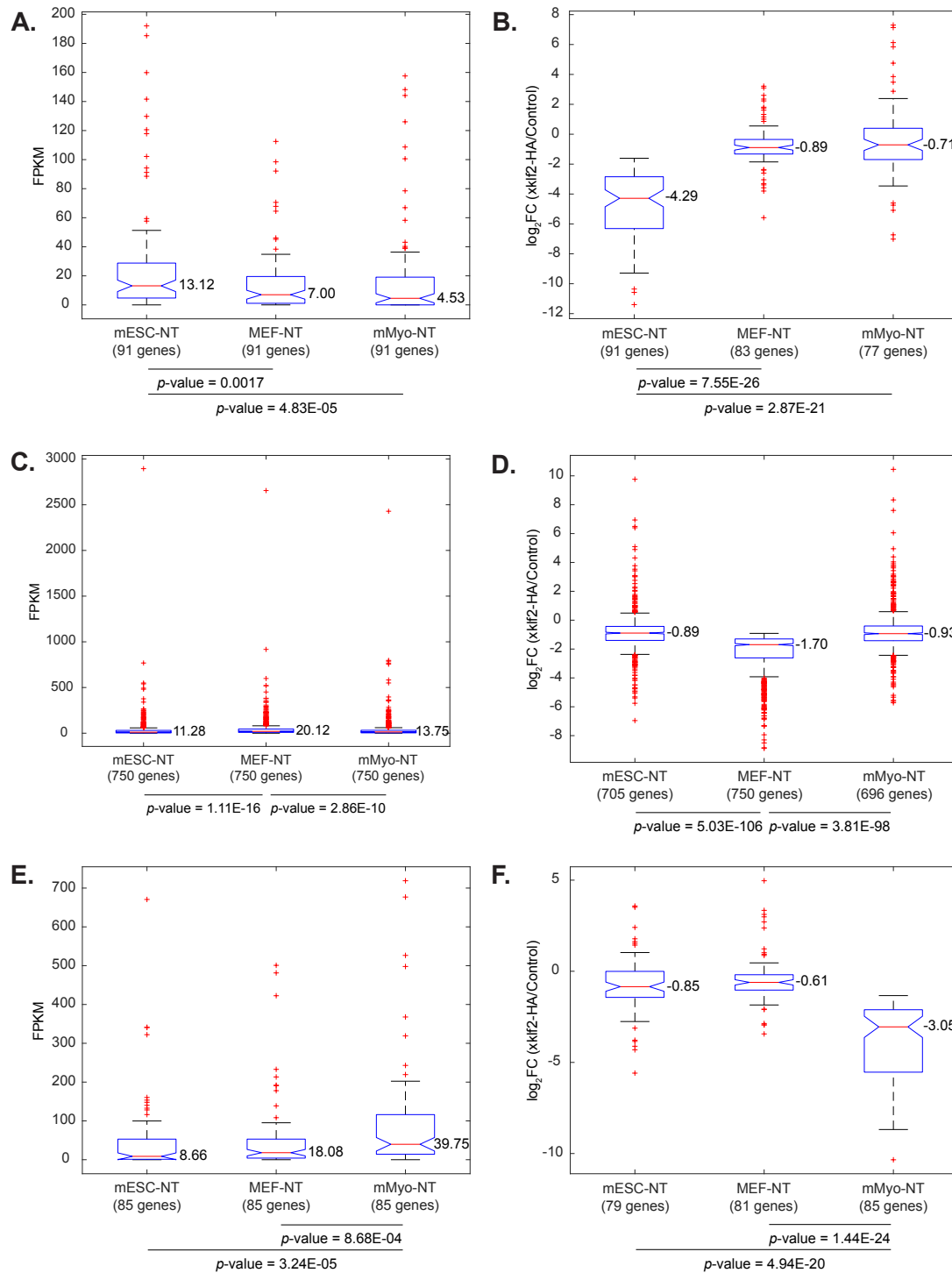


Figure 10.10.3 Down-regulated xklf2-DE genes in only one cell type have higher expression level in the absence of xklf2-HA and down-regulated by xklf2-HA overexpression with lower log₂FC than the other 2 cell types at Day 2 after Oocyte-NT.

(A, B) Down-regulated xklf2-DE genes in only mESC-NT

(C, D) Down-regulated xklf2-DE genes in only MEF-NT

(E, F) Down-regulated xklf2-DE genes in only mMyo-NT

10.11 Appendix XI: Gene ontology and KEGG pathway analysis for up-regulated and down-regulated xklf2-DE genes

To predict biological functions promoted by xklf2 if xklf2 is one of maternal factors present in *Xenopus* oocytes, up-regulated xklf2-DE genes of mESC-NT, MEF-NT and mMyo-NT are combined and applied for functional interpretation via Gene ontology enrichment analysis. 2024 in 2678 up-regulated xklf2-DE genes are input into Gene ontology database for enrichment (Biological process, Figure 10.11.1; Cellular component, Figure 10.11.2; Molecular function, Figure 10.11.3).

For GO terms of biological process, up-regulated xklf2-DE genes of mESCs, MEFs and mMyos during SCNR by oocytes are annotated mostly for developmental processes, such as animal organ morphogenesis, cell fate commitment, enzyme linked receptor protein signaling pathway, cell projection morphogenesis, cell morphogenesis involved in differentiation (terms ranked top 5 with lowest p -value) and others (all terms with p -value<0.05, Figure 10.11.1.A to E).

For GO terms of cellular component enriched, which shows the locations xklf2-DE genes perform their functions, top 5 terms enriched are synapse, synapse part, plasma membrane region, somatodendritic compartment and proteinaceous extracellular matrix (terms are ranked from lowest p -value, all terms with p -value<0.05, Figure 10.11.2). Additionally, transcription factor complex is also enriched (in red, p =0.00355, Figure 10.11.2).

For GO terms of molecular function, top 16 terms plus one term ranked 21 enriched are all about transcription, such as sequence-specific DNA binding, RNA polymerase II transcription factor activity, sequence-specific DNA binding (all terms with $p < 0.05$, Figure 10.11.3). Among them, transcription activator activity, transcription repressor activity (in red) and core promoter proximal region binding (in blue) are enriched (Figure 10.11.3). It indicates many of up-regulated *xklf2*-DE genes are transcription factors and are able to activate or repress target genes via binding on recognition sites within core promoter proximal regions.

To predict signaling pathways promoted by up-regulated *xklf2*-DE genes by *xklf2*-HA overexpression, KEGG pathway enrichment analysis is applied. 753 in 2678 up-regulated *xklf2*-DE genes of mESCs, MEFs and mMyos during SCNR by oocytes are input into KEGG pathway database for enrichment (Figure 10.11.4).

Overall, 15 pathways are enriched for up-regulated *xklf2*-DE genes ($p < 0.05$, Figure 10.11.4). Importantly, signaling pathways regulating pluripotency of stem cells (139 annotated genes in database) are enriched most significantly by up-regulated *xklf2*-DE genes and 37 up-regulated *xklf2*-DE genes are annotated in these signaling pathways ($p = 6.23 \times 10^{-7}$, Figure 10.11.4; these 37 genes are used later in Section 6.6). Therefore, *xklf2*-HA can solely promote signaling pathways regulating pluripotency of stem cells by up-regulating more than $\frac{1}{4}$ of annotated genes responsible for these pathways.

To predict functions of down-regulated *xklf2*-DE genes, GO enrichment analysis is applied. All the down-regulated *xklf2*-DE genes are combined and 809 in 1008 down-regulated *xklf2*-DE genes are input in GO database for enrichment (Biological process, Figure 10.11.5; Cellular component, Figure 10.11.6; Molecular function, Figure 10.11.7).

For enriched biological processes, DNA, RNA and protein processing related terms are enriched and top 5 terms are RNA processing, mRNA processing, mRNA metabolic process, DNA metabolic process and RNA splicing (all terms with p -value<0.05, Figure 10.11.5).

For enriched GO terms for cellular component, enriched terms are related to DNA, RNA and protein processing and top 5 terms enriched for down-regulated *xklf2*-DE genes are intracellular ribonucleoprotein complex, ribonucleoprotein complex, nuclear body, chromosome and chromosomal part (all terms with p -value<0.05, Figure 10.11.6).

For GO terms of molecular function, 6 terms are enriched for down-regulated *xklf2*-DE genes and all of them are related to DNA, RNA and protein processing, including mRNA binding, structural constituent of ribosome, helicase activity and so on (all terms with p <0.05, Figure 10.11.7).

To predict signaling pathways affected by down-regulated *xklf2*-DE genes by *xklf2*-HA overexpression, KEGG pathway enrichment analysis is applied. 283 in 1008 down-regulated *xklf2*-DE genes of mESCs, MEFs and mMyos during

SCNR by oocytes are input into KEGG pathway database for enrichment (Figure 10.11.8).

Overall, six pathways are enriched for down-regulated pathways and are related to basic cellular functions, namely DNA replication, ribosome, nucleotide excision repair, fanconi anemia pathway, cell cycle and mRNA surveillance pathway ($p < 0.05$, Figure 10.11.8). Fanconi anemia pathway is for efficiently repairing interstrand cross-links of damaged DNA.

Figure 10.11.1.A

A. Gene ontology: Biological process
Up-regulated xklf2-DE genes of all 3 cell types (List 1, Rank 1-50)

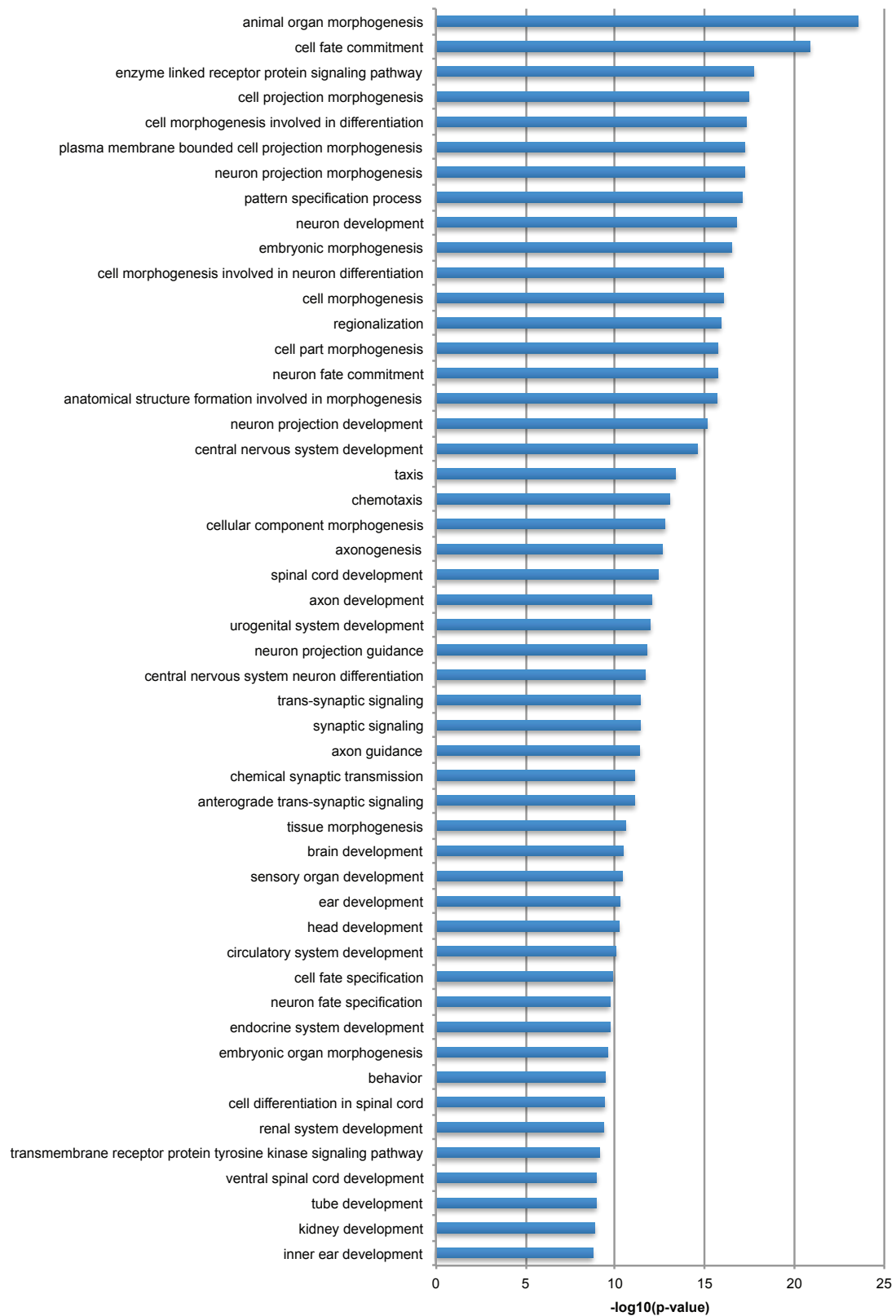


Figure 10.11.1.B

B. Gene ontology: Biological process
Up-regulated xklf2-DE genes of all 3 cell types (List 2, Rank 51-100)

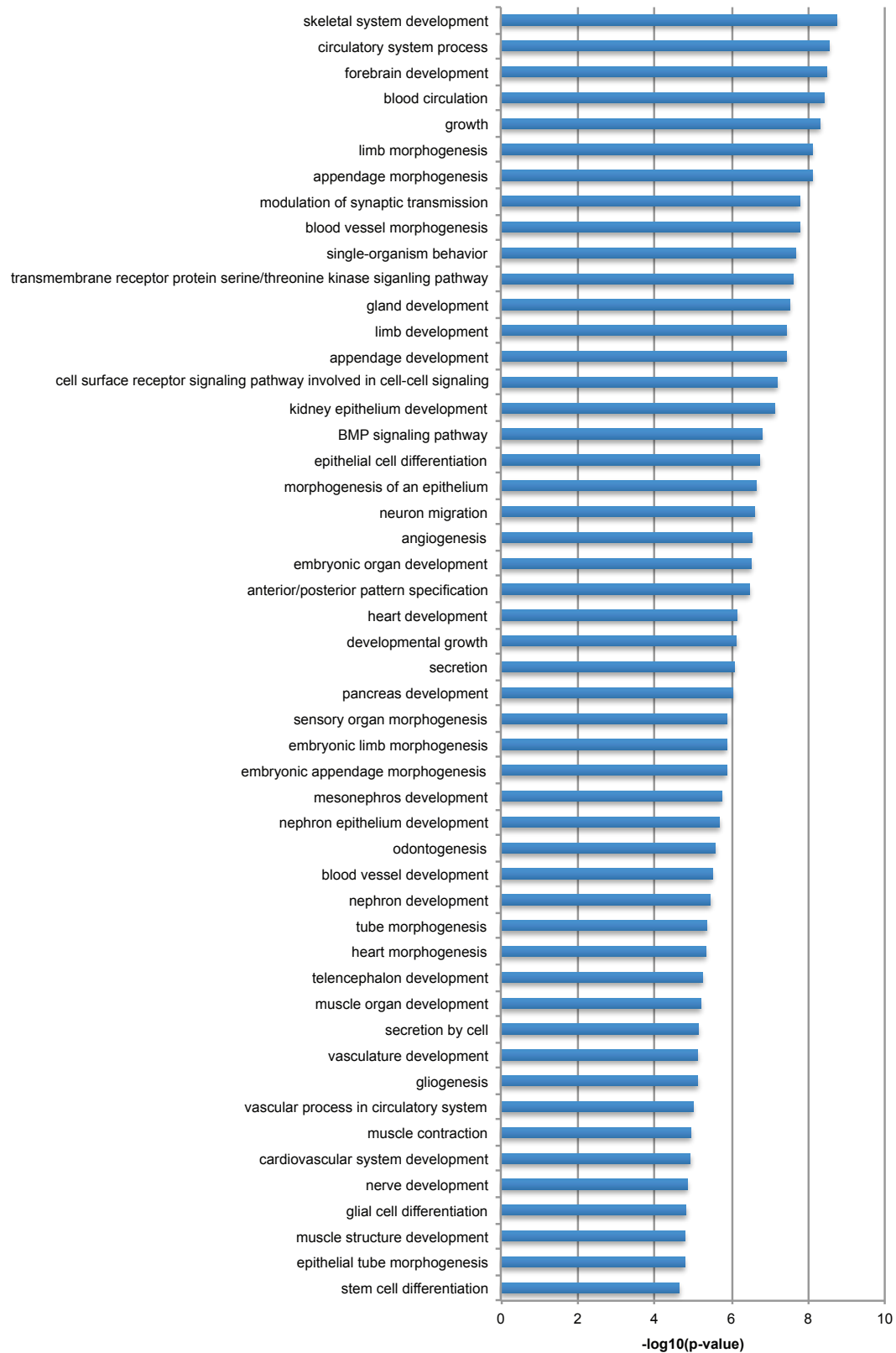


Figure 10.11.1.C

C. Gene ontology: Biological process
Up-regulated xklf2-DE genes of all 3 cell types (List 3, Rank 101-150)

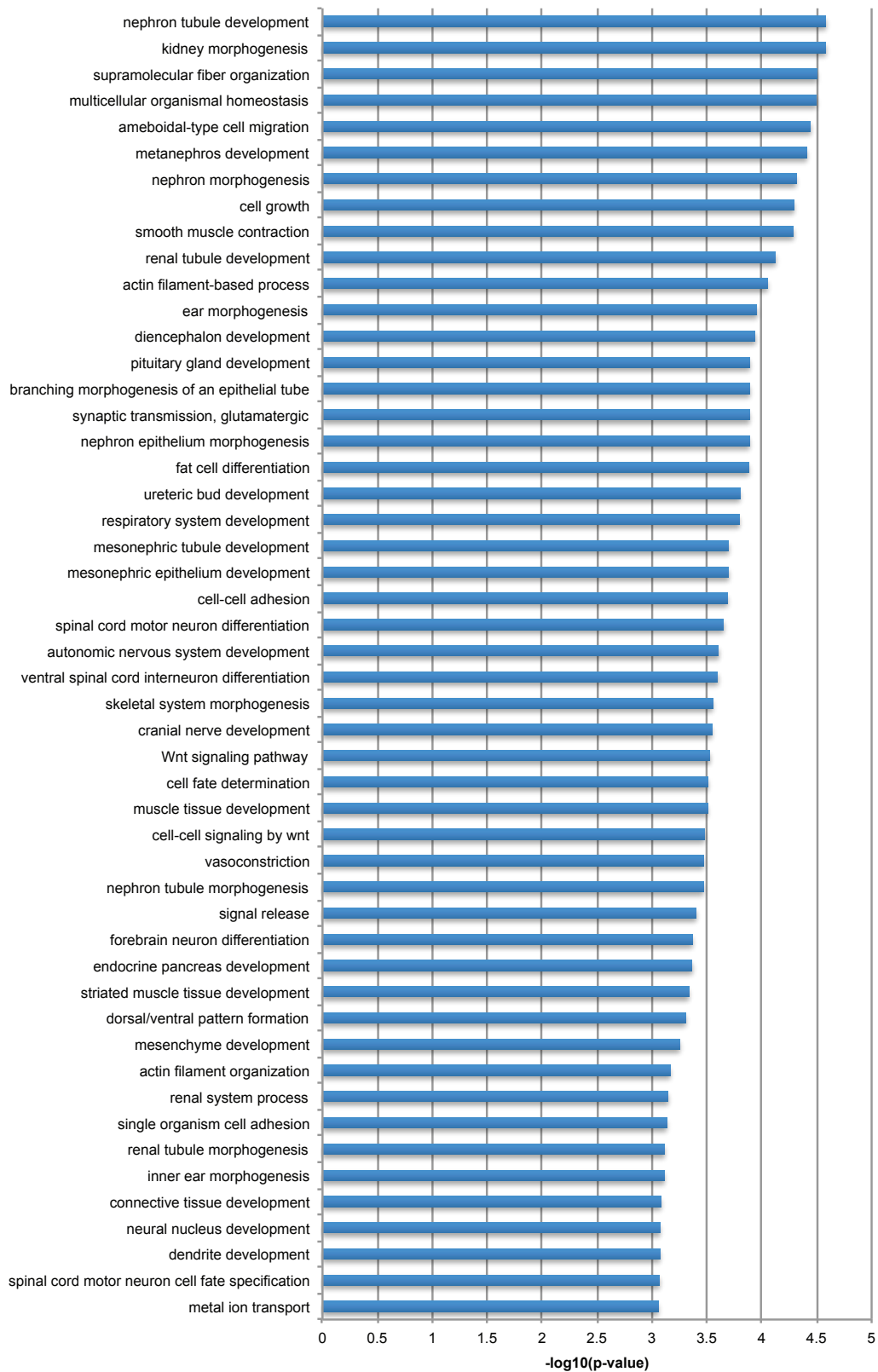


Figure 10.11.1.D

D. Gene ontology: Biological process
Up-regulated xklf2-DE genes of all 3 cell types (List 4, Rank 151-200)

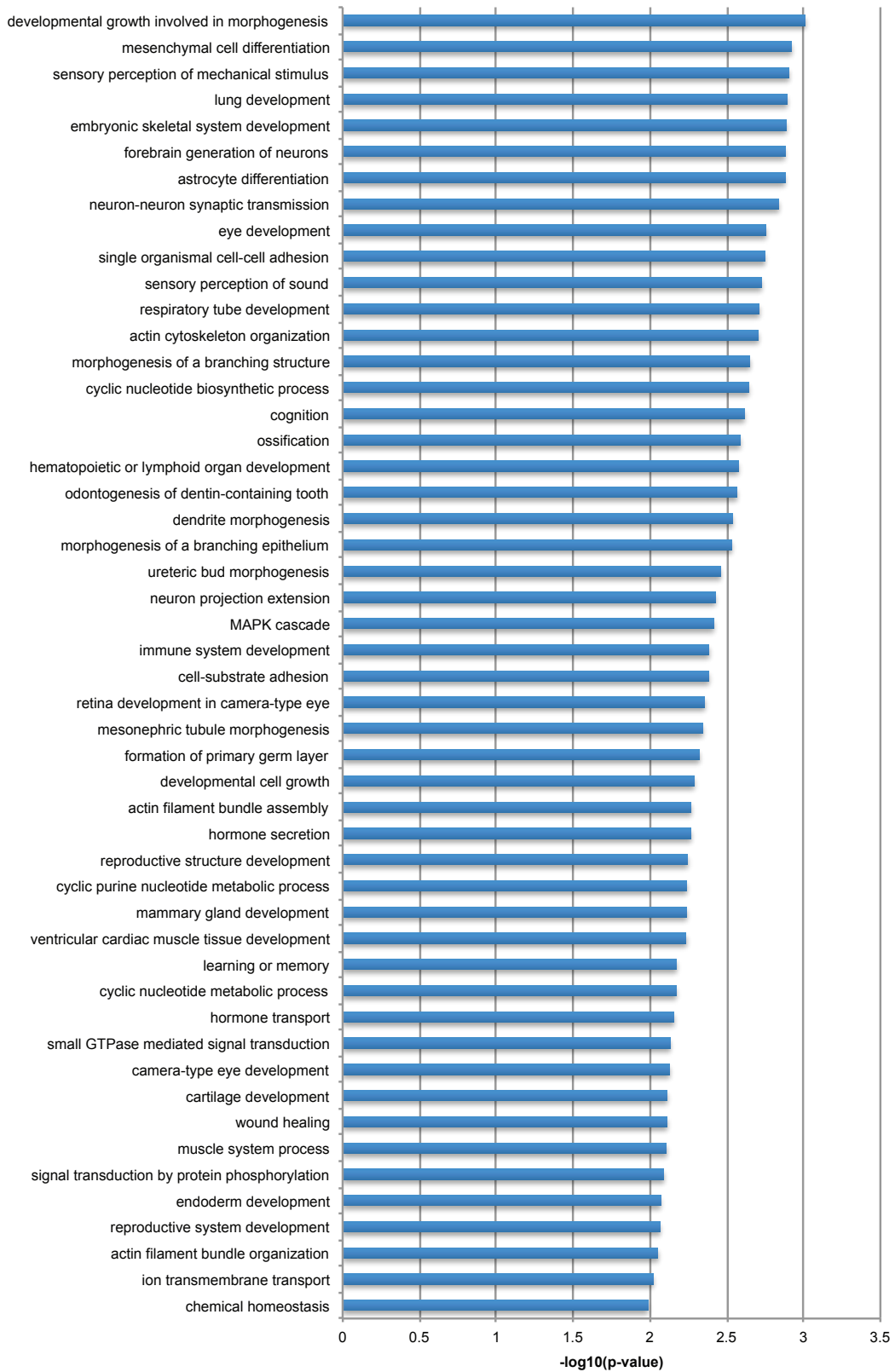


Figure 10.11.1.E

E. Gene ontology: Biological process
Up-regulated xklf2-DE genes of all 3 cell types (List 5, Rank 151-196)

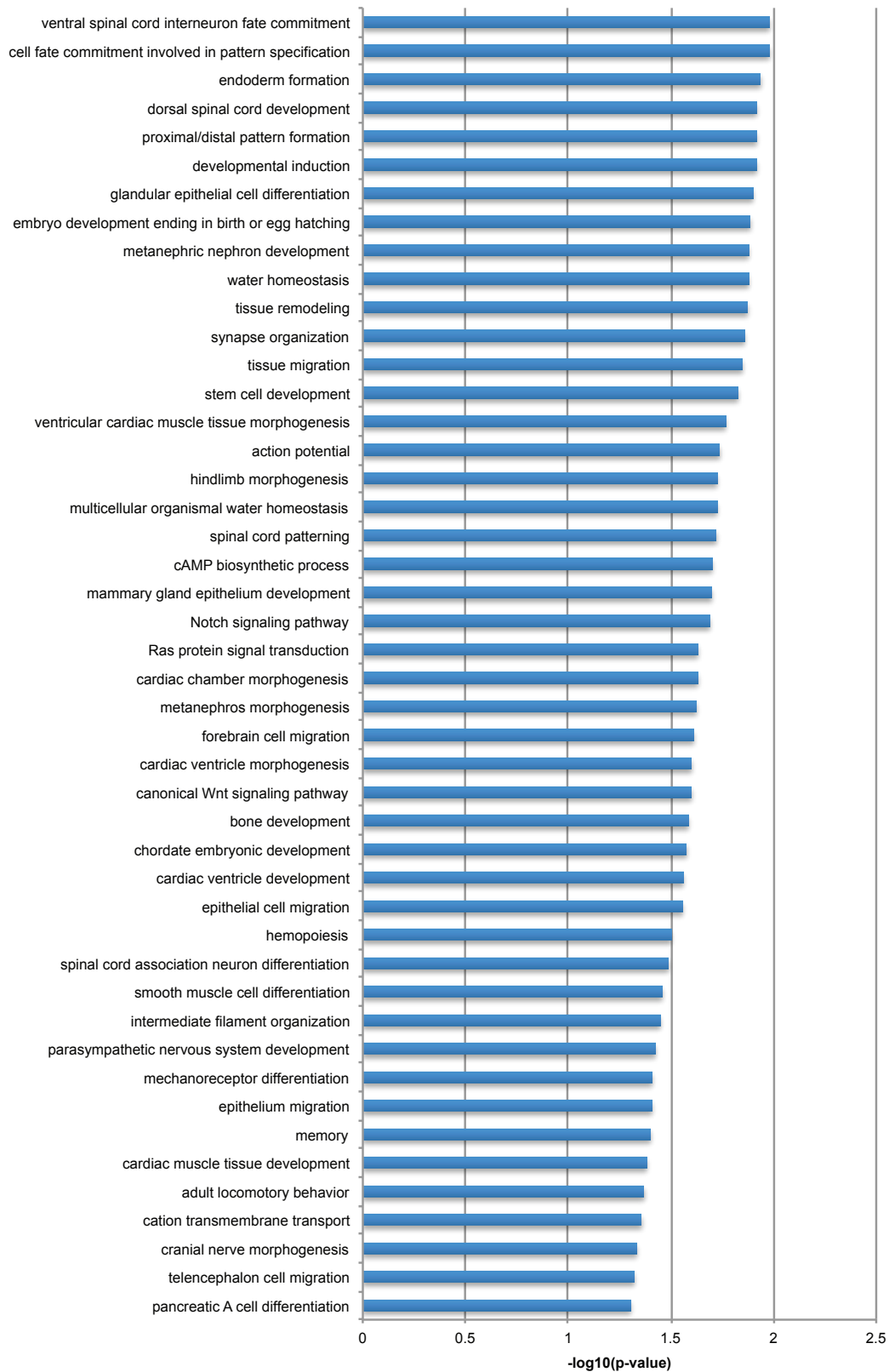


Figure 10.11.2

**Gene ontology: Cellular component
Up-regulated xklf2-DE genes of all 3 cell types**

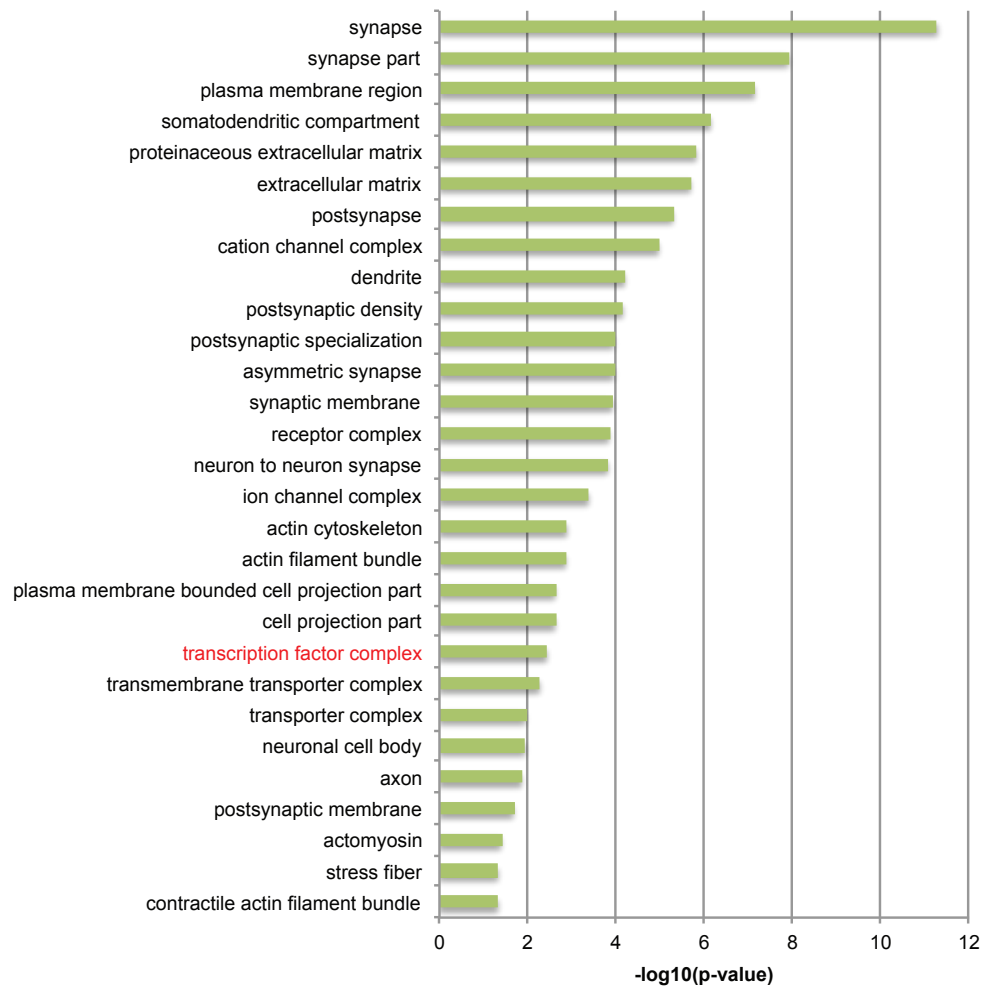


Figure 10.11.3

Gene ontology: Molecular function Up-regulated xklf2-DE genes of all 3 cell types

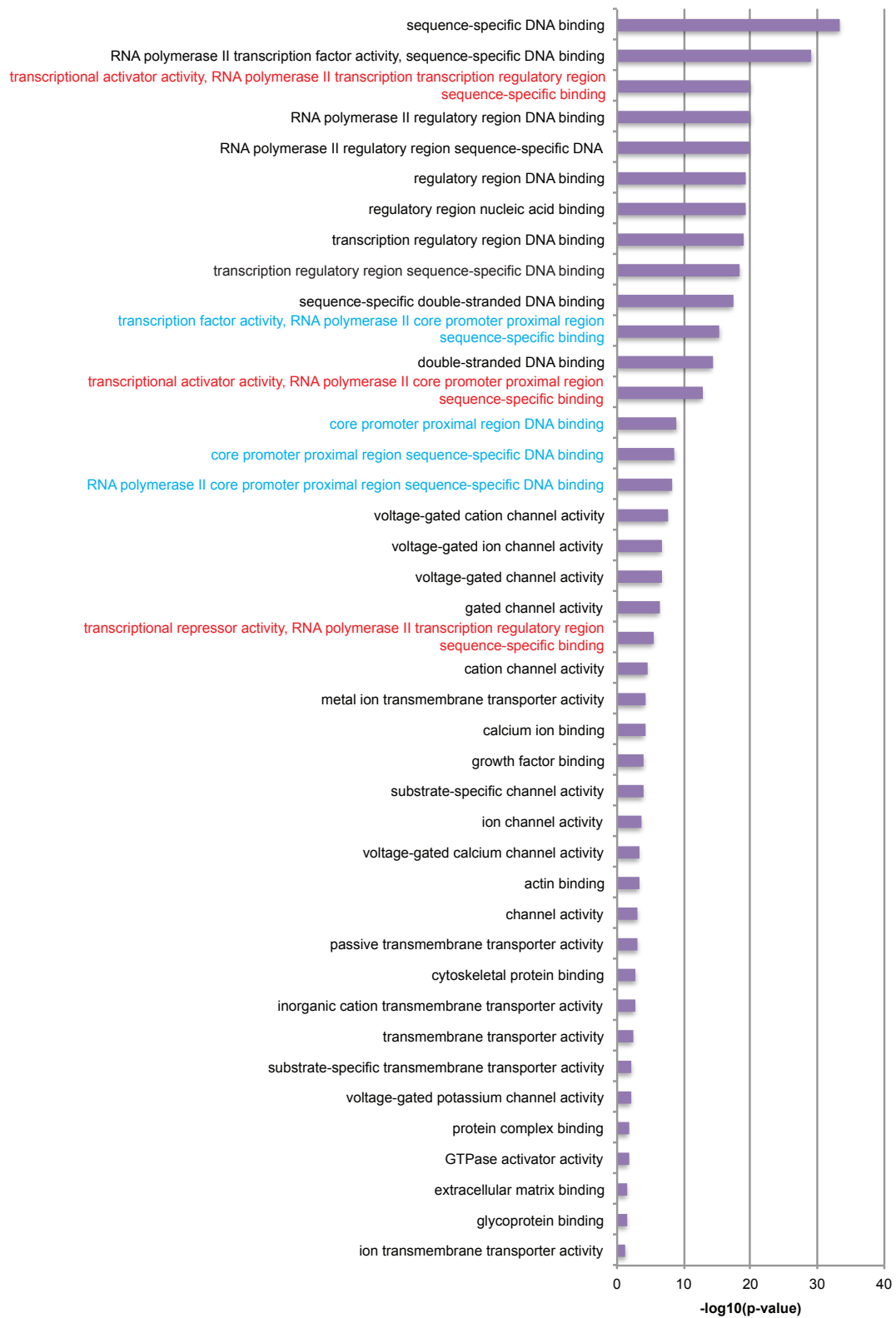


Figure 10.11.4

**KEGG pathway
Up-regulated xklf2-DE genes of all 3 cell types**

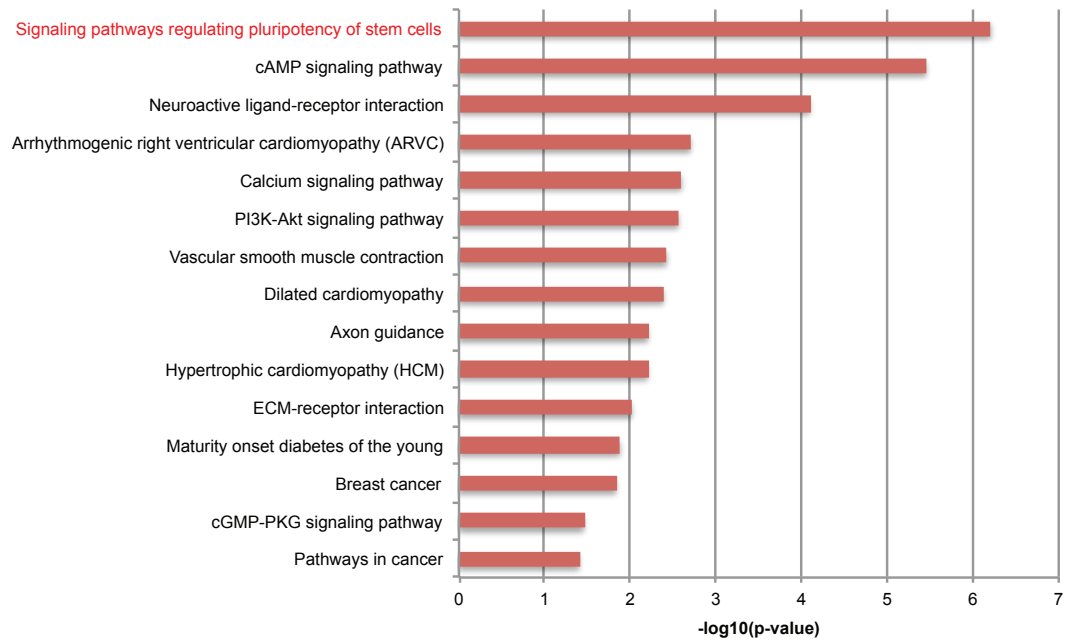


Figure 10.11.5

Gene ontology: Biological process
Down-regulated xklf2-DE genes of all 3 cell types

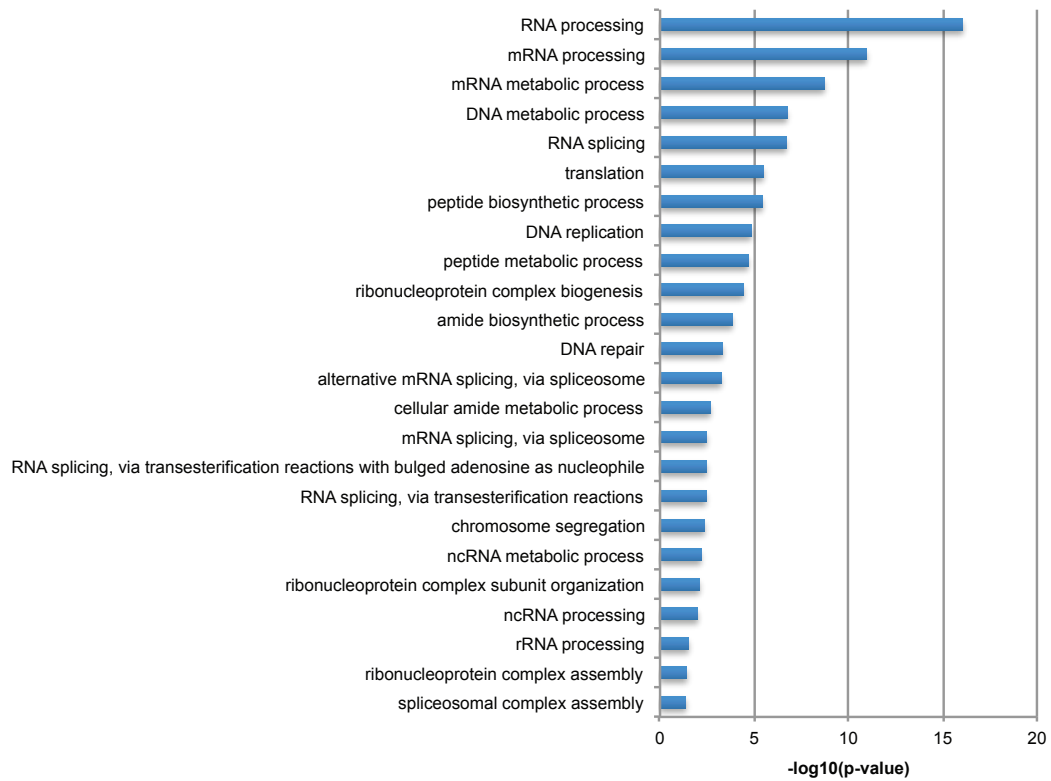


Figure 10.11.6

Gene ontology: Cellular component
Down-regulated xklf2-DE genes of all 3 cell types

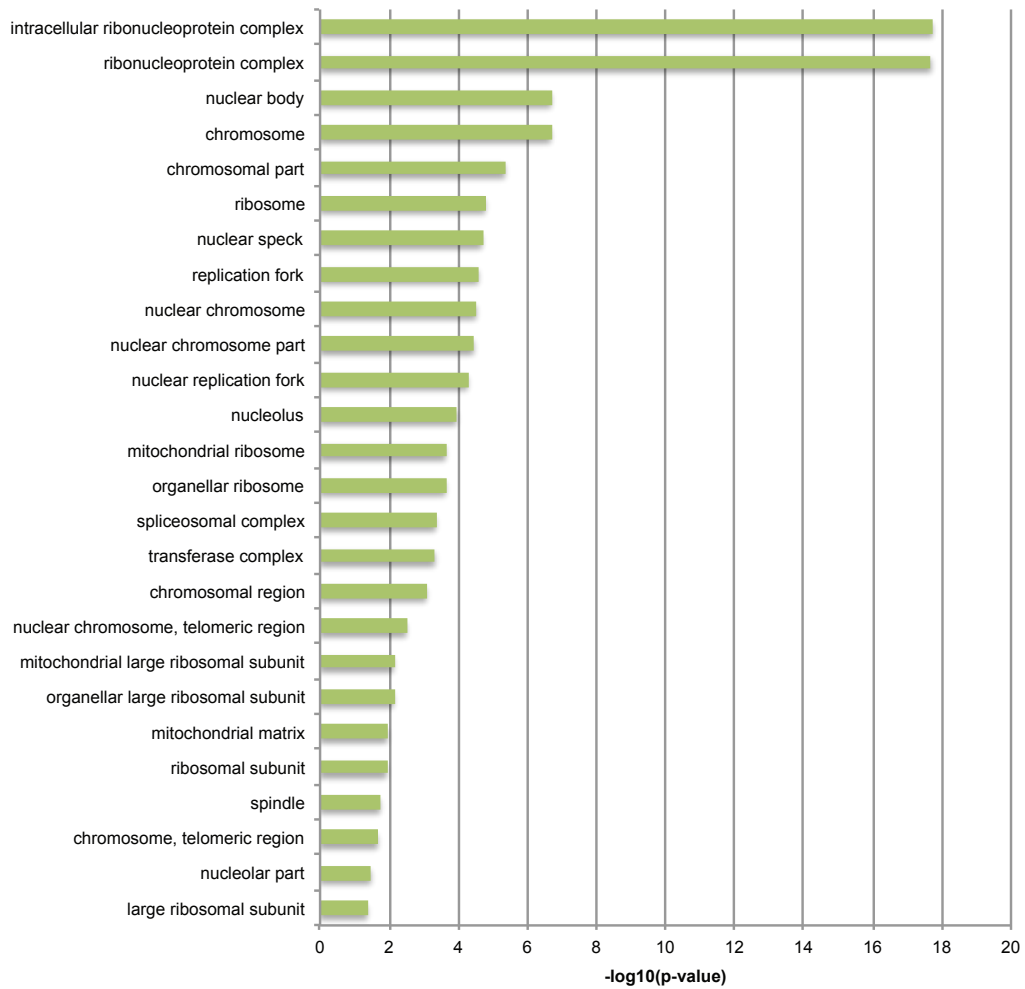


Figure 10.11.7

Gene ontology: Molecular function
Down-regulated xklf2-DE genes of all 3 cell types

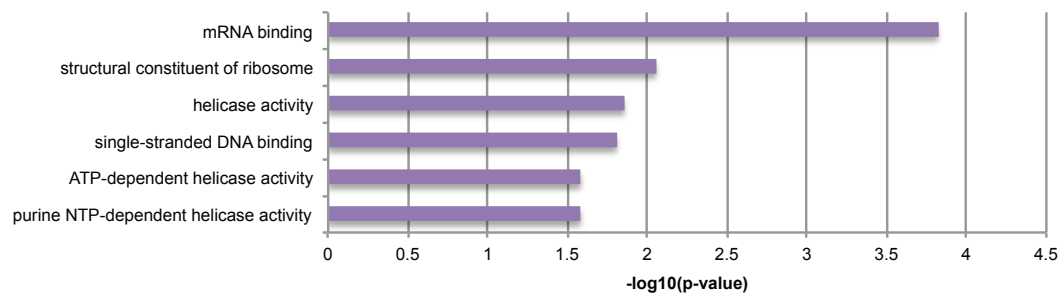
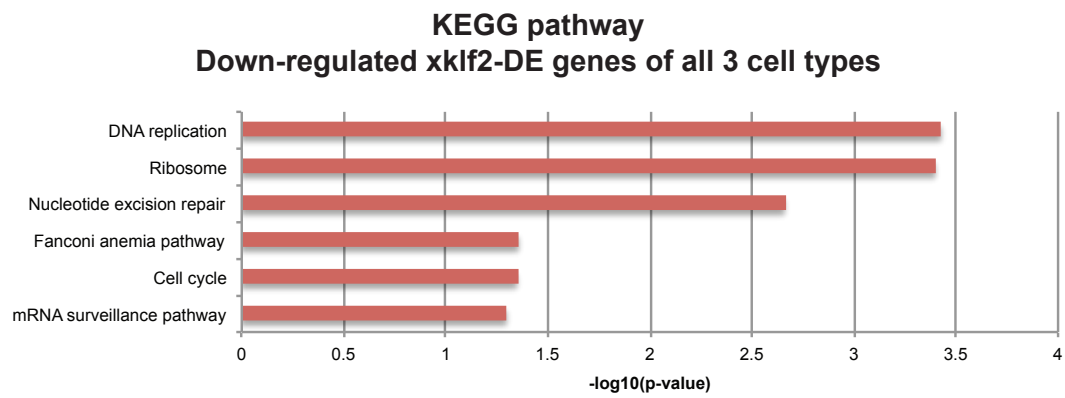


Figure 10.11.8



10.12 Appendix XII: 6.6 xklf2-HA overexpression activates resistant genes and up-regulates the expression of genes to xklf2-oocyte level

In Section 6.4, it has been shown that xklf2-HA overexpression can significantly regulate expression of genes at nil or low level to xklf2-oocyte level by comparing regulation of xklf2-DE genes in mESCs, MEFs and mMyos at Day 2 after Oocyte-NT. It has been also shown that xklf2-HA overexpression can activate SCNR resistant genes in mESC-NT, MEF-NT and mMyo-NT previously (Subsection 6.6.2, page 231).

To evaluate regulation of xklf2-DE genes by xklf2-HA overexpression after xklf2-DE genes are activated, gene expression level in mESC-NT, MEF-NT and mMyo-NT without addition of xklf2-HA is compared with log₂FC (xklf2-HA/Control groups) after the addition of xklf2-HA (Figure 10.12.1 to 10.12.6). Up-regulated xklf2-DE genes are selected for comparison because down-regulated xklf2-DE genes are usually regulated by downstream repressor of xklf2-HA (Subsection 6.2.2, page 198; 6.2.3, page 202; Appendix XI, page 369) .

For resistant genes that are silent in mESC-NT (FPKM=0) but up-regulated by xklf2-HA overexpression in mESC-NT (log₂FC>0, FDR<0.1) (Figure 6.6.2.A, page 235), the regulation of them (log₂FC, xklf2-HA/Control groups) by xklf2-HA overexpression are compared among mESC-NT, MEF-NT and mMyo-NT (Figure 10.12.1). While xklf2-DE genes of mESC-NT are silent in mESC-NT and mMyo-NT but expressed in mMyo-NT (127 genes, Figure 6.6.2.A, page

235, and 10.12.1.A), they are up-regulated by xklf2-HA overexpression significantly with more \log_2FC in mESC-NT and mMyo-NT than the same genes up-regulated by xklf2-HA overexpression in MEF-NT ($p<0.05$, mean of $\log_2FC=7.26$, 2.72 and 5.03 in mESC-NT, MEF-NT and mMyo-NT, respectively, Figure 10.12.1.B).

Likewise, when xklf2-DE genes of mESC-NT are silent in mESC-NT but expressed in MEF-NT and mMyo-NT (109 genes, Figure 6.6.2.A, page 235, and 10.12.1.C), the up-regulation by xklf2-HA overexpression in mESC-NT is significantly stronger than up-regulation by xklf2-HA overexpression in MEF-NT and mMyo-NT ($p<0.05$, mean of $\log_2FC=5.98$, 2.16 and 2.75 in mESC-NT, MEF-NT and mMyo-NT, respectively, Figure 10.12.1.D). The different extent of up-regulation by xklf2-HA overexpression can also be seen when xklf2-DE genes of mESC-NT are silent in mESC-NT and MEF-NT but expressed in mMyo-NT (38 genes, Figure 6.6.2.A, page 235; $p<0.05$, 10.12.1.E and 10.12.1.F).

For resistant genes found in MEF-NT (Figure 10.12.2) and mMyo-NT (Figure 10.12.3), it shows the different regulation of xklf2-DE genes by xklf2-HA overexpression. When the xklf2-DE genes resist activation by maternal factors, silent xklf2-DE genes (FPKM=) will be up-regulated with more \log_2FC than the same genes up-regulated by xklf2-HA overexpression in reprogrammed transcriptomes where they are expressed (FPKM>0).

For resistant genes of all 3 cell types, activation of resistant genes by xklf2-HA overexpression can also be seen (Figure 6.6.2, page 235, 10.12.4, 10.12.5 and 10.12.6). In ESC-NT, MEF-NT and mMyo-NT, 131, 307 and 153 xklf2-DE genes are silent in all 3 cell types and up-regulated by xklf2-HA overexpression (Figure 6.6.2, page 235). When comparing these resistant genes with xklf2-DE genes, which are part of reprogrammable genes and are expressed in all 3 cell types, it shows the different regulation by xklf2-HA overexpression (xklf2-DE genes in mESC-NT, Figure 10.12.4; xklf2-DE genes in MEF-NT, Figure 10.12.5; xklf2-DE genes in mMyo-NT). While reprogrammable xklf2-DE genes are expressed in all 3 cell types (FPKM>0), the up-regulation of these genes by xklf2-HA overexpression is significantly lower than up-regulation of resistant xklf2-DE genes by xklf2-HA overexpression (Figure 10.12.4, 10.12.5 and 10.12.6; $p<0.05$, Table 10.12.1, 10.12.2 and 10.12.3).

Figure 10.12.1

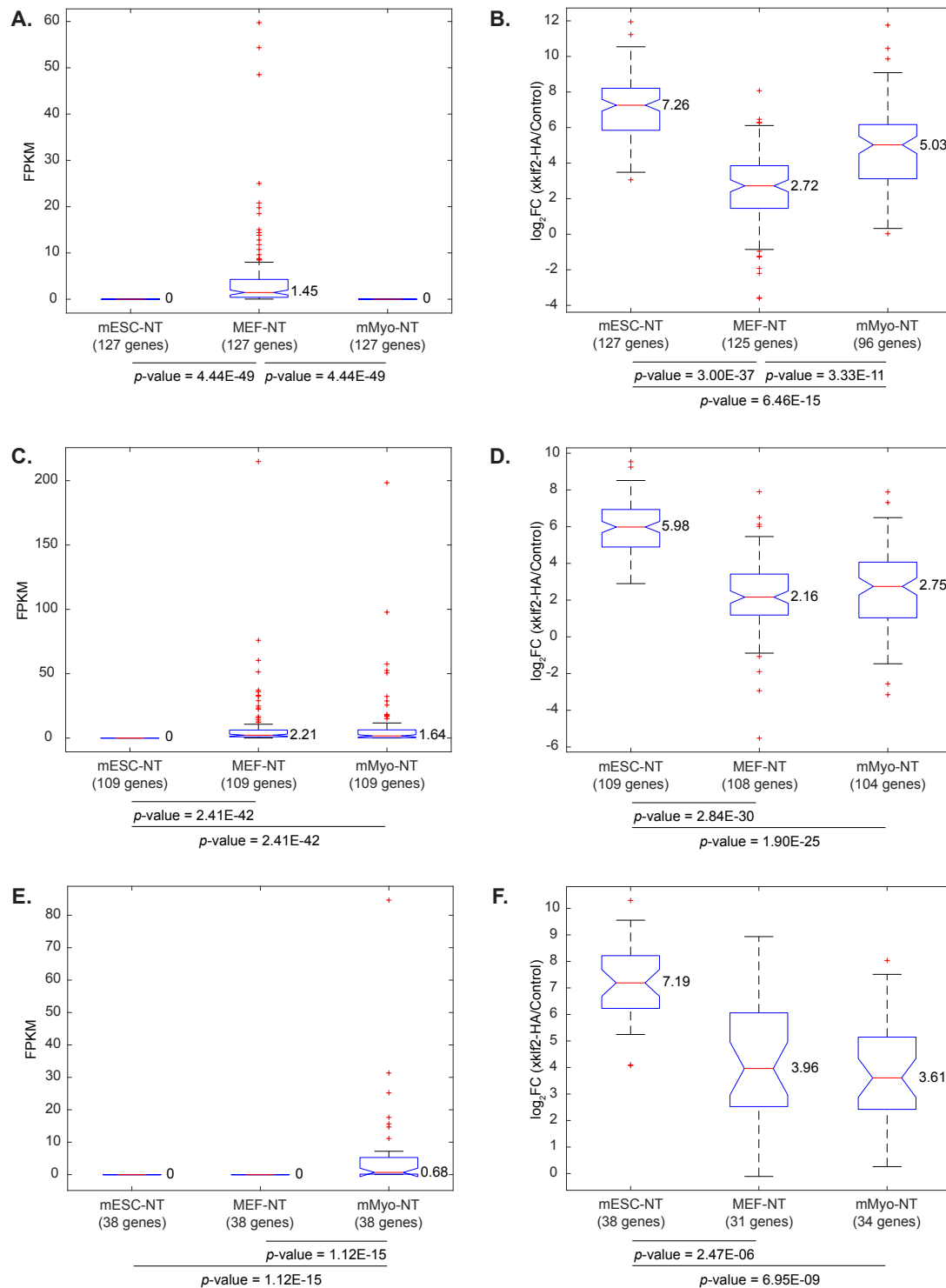


Figure 10.12.1 Up-regulated xklf2-DE genes of mESC-NT resist to be activated by maternal factors but are activated by xklf2-HA overexpression (Figure 6.6.2.A, page 235). Expression level (FPKM) and \log_2FC of SCNR resistant genes of mESC-NT expressed (A, B) in MEF-NT but not in mESC-NT and mMyo-NT, (C, D) in MEF-NT and mMyo-NT but not in mESC-NT, (E, F) in mMyo-NT but not in mESC-NT and MEF-NT is shown.

Figure 10.12.2

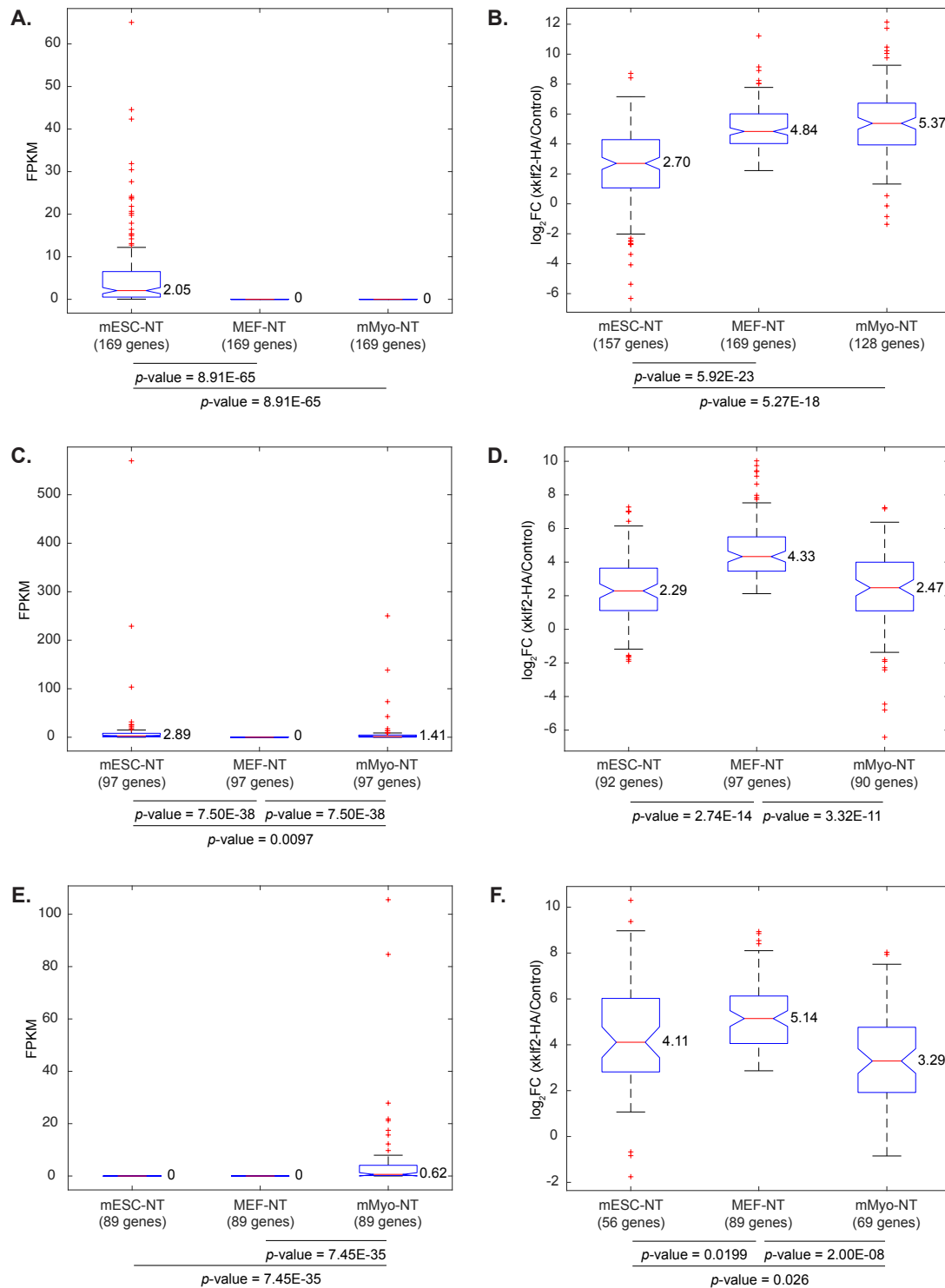


Figure 10.12.2 Up-regulated *xklf2*-DE genes of MEF-NT resist to be activated by maternal factors but are activated by *xklf2*-HA overexpression (Figure 6.6.2.B, page 235). Expression level (FPKM) and \log_2 FC of SCNR resistant genes of MEF-NT expressed (A, B) in mESC-NT but not in MEF-NT and mMyo-NT, (C, D) in mESC-NT and mMyo-NT but not in MEF-NT, (E, F) in mMyo-NT but not in mESC-NT and MEF-NT is shown.

Figure 10.12.3

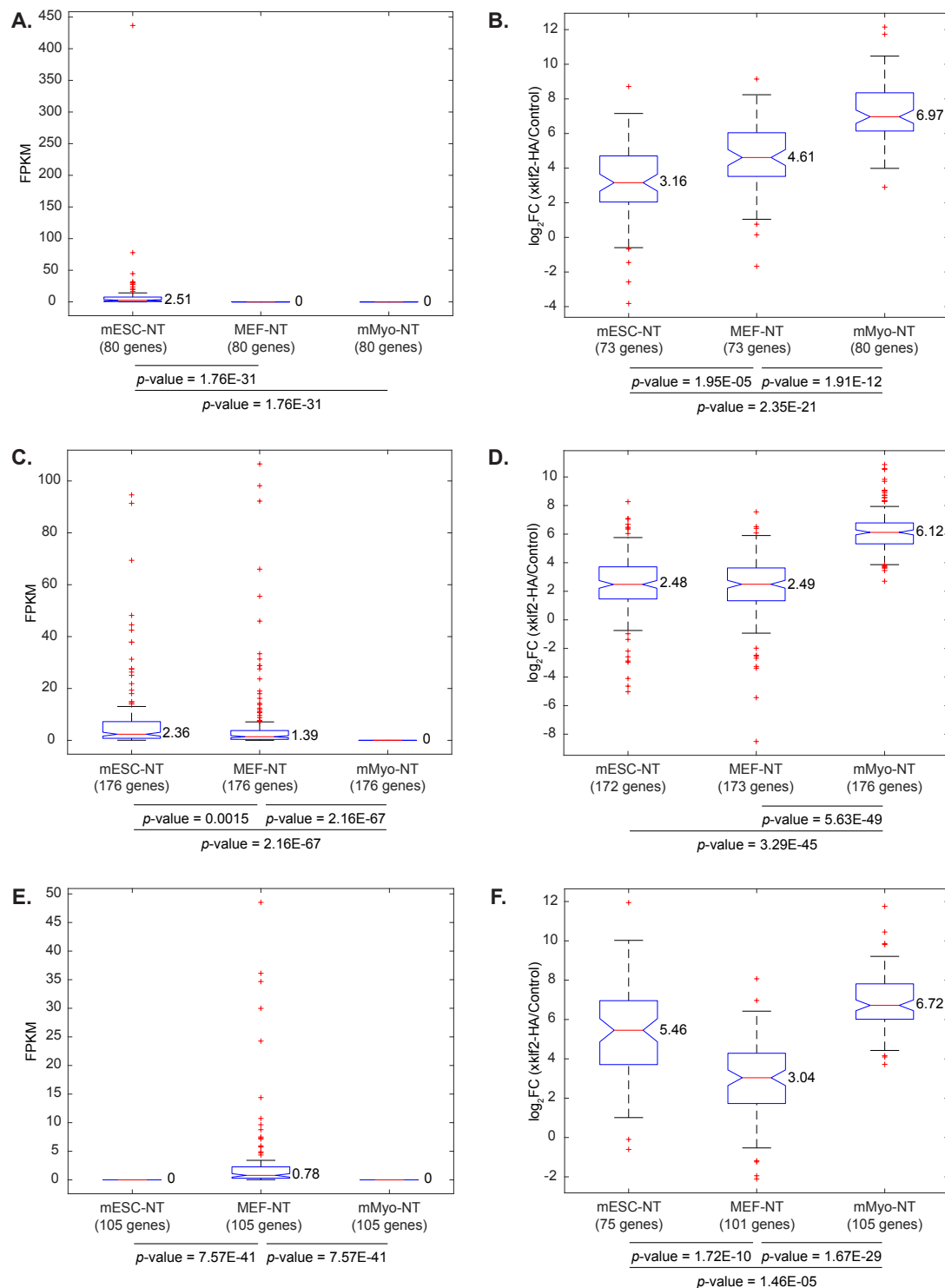


Figure 10.12.3 Up-regulated xklf2-DE genes of mMyo-NT resist to be activated by maternal factors but are activated by xklf2-HA overexpression (Figure 6.6.2.C, page 235). Expression level (FPKM) and \log_2FC of SCNR resistant genes of MEF-NT expressed (A, B) in mESC-NT but not in MEF-NT and mMyo-NT, (C, D) in mESC-NT and mMEF-NT but not in mMyo-NT, (E, F) in MEF-NT but not in mESC-NT and mMyo-NT is shown.

Figure 10.12.4

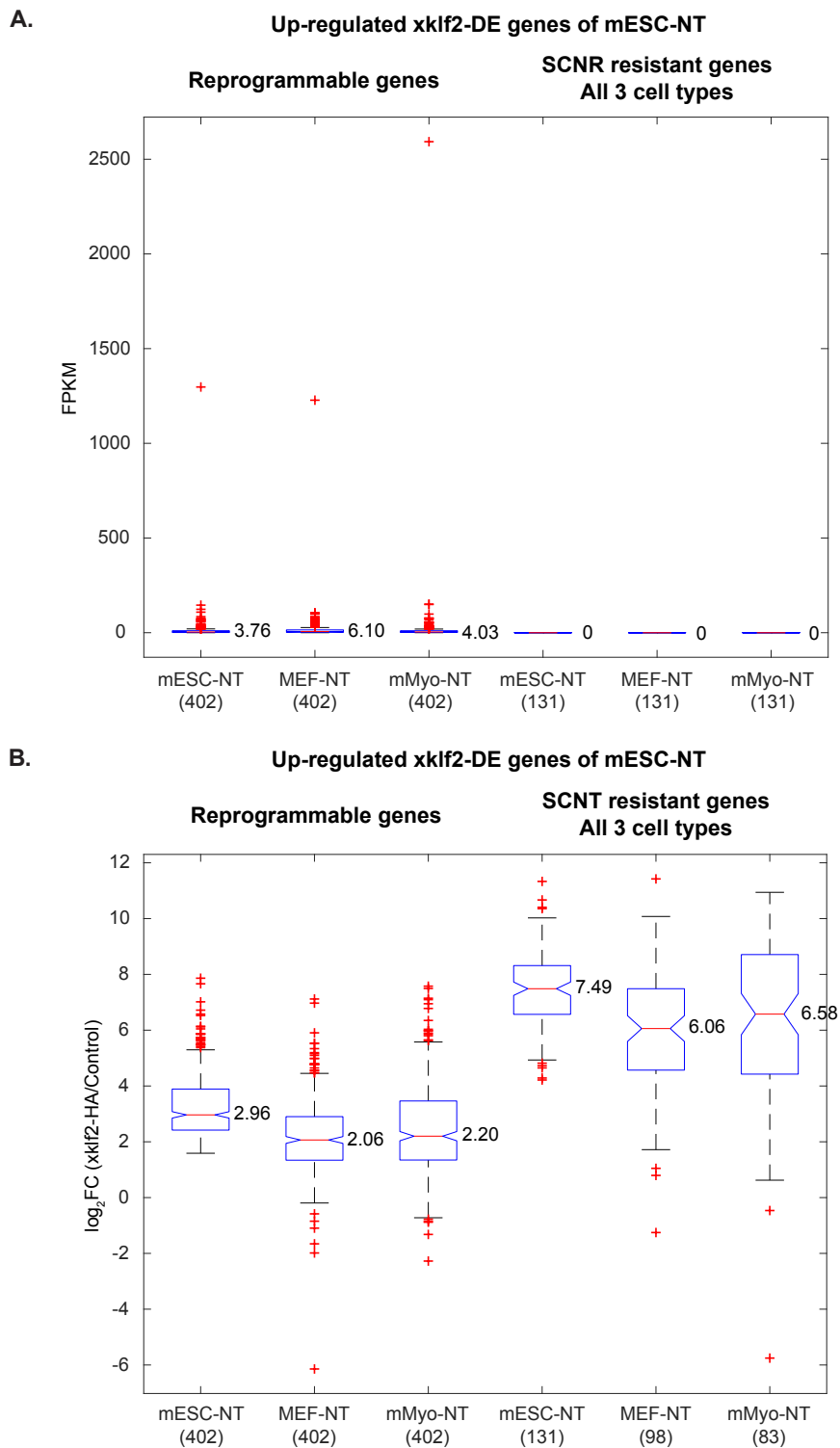


Figure 10.12.4 For up-regulated xklf2-DE genes of mESC-NT that are reprogrammable genes or SCNR resistant genes of all 3 cell types (Figure 6.6.2.A, page 235), reprogrammable xklf2-DE genes have higher expression level and are up-regulated by xklf2-HA overexpression with lower \log_2FC than SCNR resistant xklf2-DE genes significantly (p -value in Table 10.12.1).

Figure 10.12.5

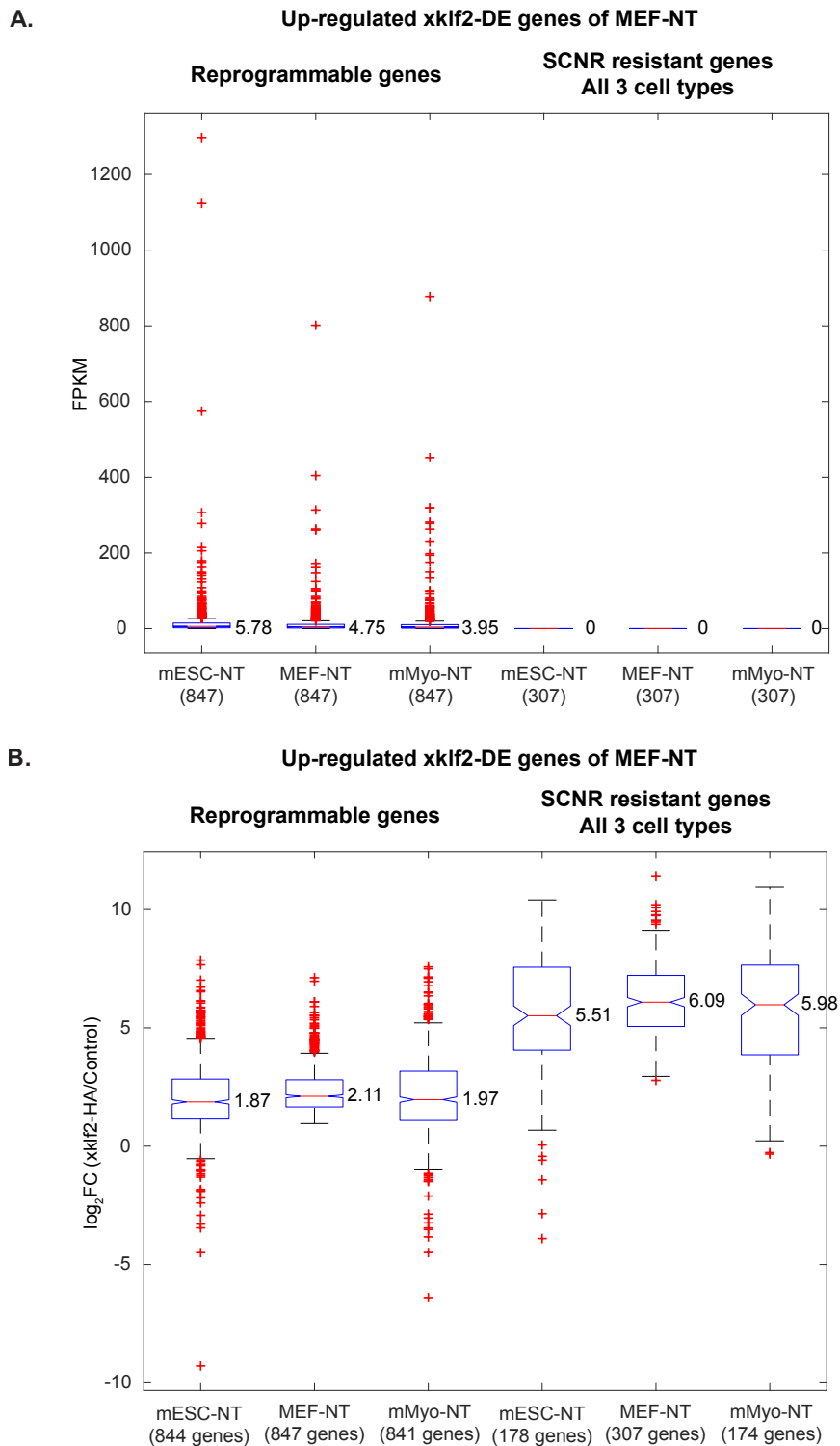


Figure 10.12.5 For up-regulated *xklf2*-DE genes of MEF-NT that are reprogrammable genes or SCNR resistant genes of all 3 cell types (Figure 6.6.2.B, page 235), different regulation by *xklf2*-HA overexpression for these two types of genes are as mentioned in Figure 10.12.4 (*p*-value in Table 10.12.2).

Figure 10.12.6

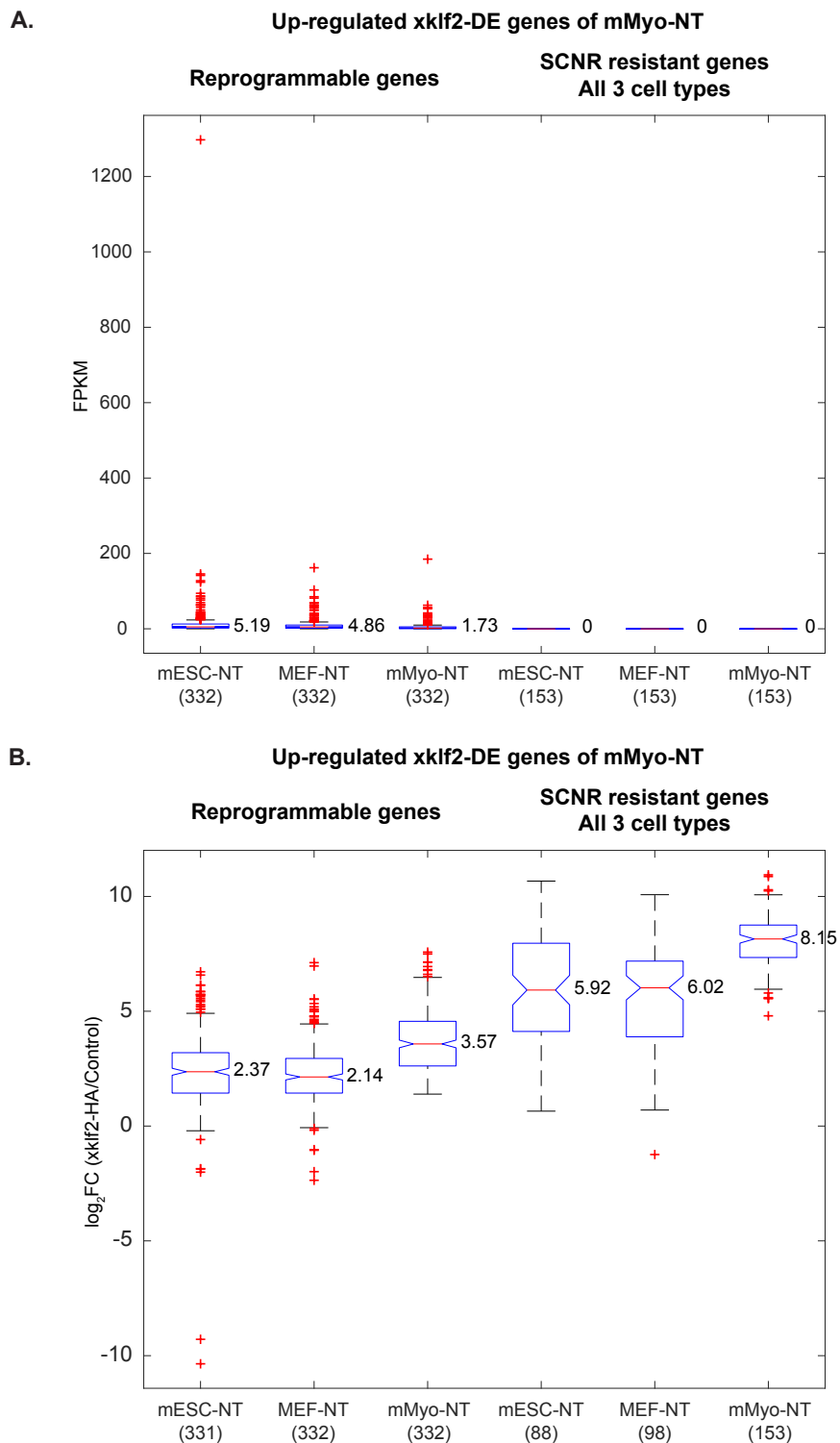


Figure 10.12.6 For up-regulated xklf2-DE genes of mMyo-NT that are reprogrammable genes or SCNR resistant genes of all 3 cell types (Figure 6.6.2.C, page 235), different regulation by xklf2-HA overexpression for these two types of genes are as mentioned in Figure 10.12.4 (p-value in Table 10.12.3).

Table 10.12.1

			FPKM						
			Reprogrammable genes of mESC			Resistant genes of mESC			
			mESC-NT	MEF-NT	mMyo-NT	mESC-NT	MEF-NT	mMyo-NT	
FPKM	Reprogrammable genes of mESC	mESC-NT							p-value
		MEF-NT	3.52E-05						
		mMyo-NT	0.7197	1.12E-04					
	Resistant genes of mESC	mESC-NT	2.83E-67	2.83E-67	2.83E-67				
		MEF-NT	2.83E-67	2.83E-67	2.83E-67				
		mMyo-NT	2.83E-67	2.83E-67	2.83E-67				

			log2FC (xklf2-HA/Control)						
			Reprogrammable genes of mESC			Resistant genes of mESC			
			mESC-NT	MEF-NT	mMyo-NT	mESC-NT	MEF-NT	mMyo-NT	
log2FC (xklf2-HA/Control)	Reprogrammable genes of mESC	mESC-NT							p-value
		MEF-NT	4.34E-34						
		mMyo-NT	2.30E-17	3.97E-02					
	Resistant genes of mESC	mESC-NT	2.62E-62	4.20E-65	8.25E-63				
		MEF-NT	3.63E-27	8.31E-38	5.46E-33	5.96E-08			
		mMyo-NT	3.67E-23	6.67E-32	2.64E-28	0.0094	0.174		

Table 10.12.1 p-value for Figure 10.12.4, p-value<0.05 is in red

Table 10.12.2

			FPKM						
			Reprogrammable genes of MEF			Resistant genes of MEF			
			mESC-NT	MEF-NT	mMyo-NT	mESC-NT	MEF-NT	mMyo-NT	
FPKM	Reprogrammable genes of MEF	mESC-NT							p-value
		MEF-NT	7.52E-04						
		mMyo-NT	1.58E-07	3.90E-02					
	Resistant genes of MEF	mESC-NT	1.02E-151	1.02E-151	1.02E-151				
		MEF-NT	1.02E-151	1.02E-151	1.02E-151				
		mMyo-NT	1.02E-151	1.02E-151	1.02E-151				

			log2FC (xklf2-HA/Control)						
			Reprogrammable genes of MEF			Resistant genes of MEF			
			mESC-NT	MEF-NT	mMyo-NT	mESC-NT	MEF-NT	mMyo-NT	
log2FC (xklf2-HA/Control)	Reprogrammable genes of MEF	mESC-NT							p-value
		MEF-NT	3.97E-09						
		mMyo-NT	0.3074	5.47E-05					
	Resistant genes of MEF	mESC-NT	1.22E-60	6.93E-62	3.69E-57				
		MEF-NT	1.19E-130	1.04E-137	1.29E-125	0.0185			
		mMyo-NT	5.74E-63	2.54E-64	1.64E-58	0.6573	0.1198		

Table 10.12.2 p-value for Figure 10.12.5, p-value<0.05 is in red

Table 10.12.3

			FPKM						
			Reprogrammable genes of mMyo			Resistant genes of mMyo			
			mESC-NT	MEF-NT	mMyo-NT	mESC-NT	MEF-NT	mMyo-NT	
FPKM	Reprogrammable genes of mMyo	mESC-NT							p-value
		MEF-NT	0.1155						
		mMyo-NT	7.92E-17	1.95E-12					
	Resistant genes of mMyo	mESC-NT	2.26E-72	2.26E-72	2.26E-72				
		MEF-NT	2.26E-72	2.26E-72	2.26E-72				
		mMyo-NT	2.26E-72	2.26E-72	2.26E-72				

			log2FC (xklf2-HA/Control)						
			Reprogrammable genes of mMyo			Resistant genes of mMyo			
			mESC-NT	MEF-NT	mMyo-NT	mESC-NT	MEF-NT	mMyo-NT	
log2FC (xklf2-HA/Control)	Reprogrammable genes of mMyo	mESC-NT							p-value
		MEF-NT	0.1283						
		mMyo-NT	8.40E-28	1.44E-38					
	Resistant genes of mMyo	mESC-NT	9.47E-30	2.43E-32	1.36E-16				
		MEF-NT	6.45E-31	1.17E-33	3.66E-17	0.3835			
		mMyo-NT	1.88E-69	1.03E-69	7.36E-68	5.16E-11	3.97E-18		

Table 10.12.3 p-value for Figure 10.12.6, p-value<0.05 is in red